THÈSE EN COTUTELLE PRÉSENTÉE

POUR OBTENIR LE GRADE DE

DOCTEUR DE

L'UNIVERSITÉ DE BORDEAUX

ET DE L'UNIVERSITÉ DE NAPLES FEDERICO II

SCIENCES CHIMIQUES

SPÉCIALITÉ CHIMIE ORGANIQUE

BIOTECHNOLOGIE

Par Marian VINCENZI

TARGETING PROTEIN INTERACTIONS WITH BIOTECHNOLOGICAL ORIGINAL MOLECULES: A NMR AND MOLECULAR MODELLING INTEGRATED APPROACH

Sous la direction de Jean GUILLON
et de Filomena ROSSI

Soutenue le 22 Avril 2016

Membres du jury :

M. LISOWSKI, Vincent	Professeur Université de Montpellier	Président
M. LISOWSKI, Vincent	Professeur Université de Montpellier	Rapporteur
Mme SAVIANO, Gabriella	Professeur Université de Molise	Rapporteur
M. GUILLON, Jean	Professeur Université de Bordeaux	Examinateur
Mme ROSSI, Filomena	Professeur Università degli Studi di Napoli	Examinateur
Mme RONGA, Luisa	Maître de conférences Université de Bordeaux	invité
Mme LEONE, Marilisa	Chercheur Institute of Biostructures and Bioimaging	invité

INTERACTIONS PROTÉINES-MOLÉCULES BIOTECHNOLOGIQUES ORIGINALES: UNE APPROCHE INTÉGRÉE DE RMN ET DE MODÉLISATION MOLÉCULAIRE

Résumé :

Ce travail de thèse PhD concerne l'application d'une approche intégrée pour obtenir une meilleure compréhension des mécanismes d'action de Akt et CXCR4, surexprimées dans différents cancers humains. Efforts récents dans le développement et l'évaluation biologique (activité antiprolifération) de petites entités moléculaires inhibitrices d'Akt, une sérine/thréonine protéine-kinase, ont conduit à l'identification de nouveaux inhibiteurs pyrrolo[1,2-*a*]quinoxaline, conçus et préparés via une stratégie multi-étapes. Certains des composés synthétisés ont montré une activité contre les lignées cellulaires leucémiques testées (Jurkat, U266, K562, U937 et HL60) supérieure à celle de composé de référence A6730. En outre, des résultats préliminaires menés sur Akt puis sur le domaine de PH isolé d'Akt, ont montré que ils peuvent être considérés comme des inhibiteurs allostériques potentiels. La seconde partie des travaux concerne la conception et la synthèse de deux nouvelles séquences peptidiques contenant quelques acides aminés "disorder promoting" et une unité CPC. Les études CD, RMN et MD ont fait ressortir leur flexibilité et ont démontré leurs capacités à assumer des ensembles de conformations stabilisées par un réseau de liaisons hydrogène. Ensuite, nous avons étudié l'effet de la chaîne alkyle reliée sur la conformation des peptides. Des études de fluorescence et DLS ont été réalisées pour évaluer les CMC et la dimension des agrégats supramoléculaires. Les tests biologiques ont souligné que ces édifices moléculaires (peptides amphiphiles nommés, PAs) montrent ainsi des activités prometteuses, voire plus que la molécule de référence (AMD3100).

Mots clés :

Akt; pyrrolo[1,2-*a*]quinoxaline; allostérie ; CXCR4 ; CXCL12 ; CPC ; amphiphiles

TARGETING PROTEIN INTERACTIONS WITH BIOTECHNOLOGICAL ORIGINAL MOLECULES: A NMR AND MOLECULAR MODELLING INTEGRATED APPROACH

Abstract :

This PhD thesis work has been covered in the application of an integrated approach to get a better understanding about the mechanism of action of two systems: Akt and CXCR4, proteins overexpressed in different human cancers. On the basis of previous results obtained on the antiproliferative activities of small molecule inhibitors of Akt, a serine/threonine protein kinase, a novel series of pyrrolo[1,2-a]quinoxaline derivatives have been designed and synthesized via multistep heterocyclization process. Some compounds showed promising activities against all leukemia cell lines tested (Jurkat, U266, K562, U937 and HL60), even better than the reference compound (A6730) one. In addition, docking results, conducted on the isolated PH domain, showed that these new compounds could be considered as allosteric inhibitors. The second workpackage reports on the design and the synthesis of two new peptidic sequences containing a few amino acids “disorder promoting” and a CPC unit, the CXCL12 binding motif towards CXCR4. The peptide structural preferences were analysed by CD, NMR and MD techniques that highlighted their flexibility and demonstrated the ability of these peptides to assume conformational ensembles stabilized by a network of transient and dynamic H-bonds. Afterwards we studied the alkyl chain effect on the conformation of the peptide portion. Solution fluorescence and DLS studies have been performed to evaluate CMC and the dimension of supramolecular aggregates (named peptide amphiphiles, PAs). Biological tests pointed out that these molecular buildings show promising activities, even higher than the reference molecule (AMD3100) one.

Keywords :

Akt; pyrrolo[1,2-*a*]quinoxaline; allostery ; CXCR4 ; CXCL12 ; CPC ; amphiphiles

---

Unité de recherche

[INSERM U1212 / UMR CNRS 5320; Université de Bordeaux, 146, rue Léo Saignat, 33076 Bordeaux cedex, France]

TESI IN CO-TUTELA ITALIA-FRANCIA

TARGETING PROTEIN INTERACTIONS WITH BIOTECHNOLOGICAL ORIGINAL MOLECULES: A NMR AND MOLECULAR MODELLING INTEGRATED APPROACH

MARIAN VINCENZI

École Doctorale des Sciences Chimiques
Spécialité Chimie Organique
Université de Bordeaux

Dottorato in Scienze Biotecnologiche- XXVIII Ciclo
Indirizzo Biotecnologie Molecolari ed Industriali
Università degli Studi di Napoli Federico II

Dottorato in Scienze Biotecnologiche- XXVIII Ciclo
Indirizzo Biotecnologie Molecolari ed Industriali
Università degli Studi di Napoli Federico II

TARGETING PROTEIN INTERACTIONS WITH BIOTECHNOLOGICAL ORIGINAL MOLECULES: A NMR AND MOLECULAR MODELLING INTEGRATED APPROACH

DOTTORANDO: Marian Vincenzi

RELATORI: Prof.ssa Filomena Rossi

Prof. Jean Guillon

COORDINATORE: Prof. Giovanni Sannia

INDEX

ABBREVIATIONS

Amino Acids: singol and 3 letters code

K	Lys	Lysine
P	Pro	Proline
C	Cys	Cysteine
G	Gly	Glycine
E	Glu	Glutamic Acid
Y	Tyr	Tyrosine

2-Cl-(Trt)-Cl	2-Chlorotrityl chloride resin
AcOEt	Ethyl acetate
AhOh	21-Amino-4,7,10,13,16,19-hexaoxaheneicosanoic acid
Akt	Protein kinase B
ANS	8-Anilino-1-naphthalenesulfonic acid ammonium salt
cAMP	Cyclic Adenosine Monophosphate
CAPITO	Cd Analysis & Plotting TOol
CARA	Computer Aided Resonance Assignment
CCRF–CEM	Human drug sensitive cancer cell lines
CD	Circular Dichroism
CMC	Critical Micellar Concentration
CXCL12	natural ligand of CXCR4
CXCR4	C-X-C chemokine receptor type 4
D	Diffusion coefficients
DABCO	1,4-diazabicyclo[2.2.2]octane
DBU	1,8-diazabicyclo[5.4.0]undec-7-ene
DCM	Dichloromethane
DIEA	N,N-diisopropylethylamine
DLS	Dynamic Light Scattering
DMF	Dimethylformamide
DPFGSE	Double Pulsed Field Gradient Selective Echo
dppp	1,3-bis(diphenylphosphino)propane
DQFCOSY	Double Quantum Filter Correlation Spectroscopy
DSS	Dihedral Space Sampling
ESI	Electro Spray Ionozation
EtOH	Ethanol
Et2O	Ethyl ether
FCS	Fetal Calf Serum
Fmoc	Fluorenylmethyloxycarbonyl
HBTU	2-(1H-Benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate
HRMS	High Resolution Mass Spectrometry
IDP	Intrinsically Disordered Peptide
IR	Infrared
IS	Selectivity Index
LC	Liquid Chromatography
MD	Molecular Dynamic
MeOH	Methanol

MS	Mass Spectroscopy
MTS	3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide
NMR	Nuclear Magnetic Resonance
NMP	*N*-methyl-2-pyrrolidone
NOESY	Nuclear Overhauser Enhancement Spectroscopy
PA	Peptide Amphiphiles
PBMNC	Peripheral Blood Mononuclear Cells
PDI	Polydispersity Indices
PEG	Polyethylene Glycol
PHA	Polyhydroxyalkanoates
pip	Piperidine
PTSA	*p*-Toluene Sulfonic Acid
pyr	Pyridine
QSAR	Quantitative Structure Activity Relationship
RH	Hydrodynamic Radii
RMSD	Root Mean Square Deviation
RMSF	Root Mean Square Fluctuation
ROESY	Rotating frame Overhauser Enhancement SpectroscopY
RP-HPLC	Reverse Phase-High Performance Liquid Chromatography
SAR	Structure Activity Relationship
SD	Deviation Standard
THF	Tetrahydrofuran
TIS	Tri-Isopropyl-Silane
TFA	Trifluoroacetic Acid
TOCSY	TOtal Correlation SpectroscopY
Trt	Trityl
TSP	TrimethylSilyl-3-Propionic acid sodium salt-d4

Abstract
This PhD thesis work has been covered in the application of an integrated approach of different experimental and computational techniques, to get a better understanding about the solution structure details and the intermolecular interactions of two systems: Akt-heterocyclic ligand (**Workpackage 1**) and the chemokine, CXCR4 (**Workpackage 2**). It is well known that the dysregulation of Akt and CXCR4 proteins, normally involved in regulation processes in cell, such as proliferation, is overexpressed in different human cancers. For this reason, they are considered as valuable targets in the development of potential cancer therapeutic agents. Recent efforts in the development and biological evaluation of small molecule inhibitors of Akt, a serine/threonine protein kinase, have led to the identification of novel inhibitors with various heterocycle scaffolds. Among heterocyclic compounds that attracted a lot of attention because of its wide spread biological activities, the pyrrolo[1,2-*a*]quinoxaline heterocyclic framework has been identified as interesting scaffolds for antiproliferative activity against various human cancer cell lines. Thus, based on previous results obtained on the antiproliferative activities of new pyrrolo[1,2-*a*]quinoxalines, a novel series of pyrrolo[1,2-*a*]quinoxaline derivatives **1a-l** have been designed and synthesized *via* multistep heterocyclization process. Their cytotoxicities were evaluated against five different leukemia cell lines, including Jurkat and U266 (lymphoid cell lines), and K562, U937, HL60 (myeloid cell lines), as well as normal human peripheral blood mononuclear cells. The new series showed promising cytotoxic potential against all leukemia cell lines tested, and some compounds showed better results than the reference compound A6730. Some compounds, such as **1a**, **1e**, **1g** and **1h** are promising because of their high activity against leukemia and their low activity against normal hematopoietic cells. Structure-activity relationships of these new synthetic compounds **1a-l** are here also discussed. In addition, docking results, conducted on the isolated PH domain, showed that these new compounds could be considered as allosteric inhibitors.
The second workpackage reports on the design and the synthesis of two new peptidic sequences containing a few amino acids with disorder propensity. In addition, they present a CPC unit: binding motif of the specific ligand of CXCR4 namely the CXCL12. The peptide structural preferences were analysed by CD and NMR techniques that highlighted their flexibility. These results were also confirmed by MD simulations that demonstrated the ability of these peptides to assume conformational ensembles stabilized by a network of transient and dynamic H-bonds and characterized by interactions with water molecules. Afterwards we have explored the possibility that these peptides could be used as polar heads connected to alkyl chains to generate new molecular buildings (named peptide amphiphiles, PAs) for drug delivery vehicles in cells over-expressing the CXCR4 receptor. For these reasons, we firstly studied the alkyl chain effect on the conformational rearrangement of intrinsically disordered peptides. Solution fluorescence and DLS studies have been performed to evaluate critical micellar concentrations and the dimension of supramolecular aggregates. Biological tests pointed out that a hydrophobic moiety connected to our peptides could improve their activity on CXR4. In fact, these molecular buildings show promising activities, even higher than the reference molecule (AMD3100). According to these preliminary results, these PA systems could be further modified to realize supramolecular peptide-based compounds, able to be encapsulated by living cells for biomedical application.

Riassunto

Il processo di ricerca e sviluppo di nuovi farmaci, come è noto, richiede numerose fasi. Queste includono, oltre all' identificazione del *target* biologico, la sintesi, lo *screening* e la successiva costruzione di biblioteche di molecole potenzialmente dotate di attività farmacologica. L'intero processo è molto impegnativo dal punto di vista dei costi (1-2 miliardi di euro) e dei tempi (12-13 anni) [1]. Pertanto, i metodi della chimica computazionale e la dinamica molecolare hanno le potenzialità di contribuire allo sviluppo di un farmaco in termini di progettazione, razionalizzazione delle relazioni struttura-attività e quindi di riduzione dei costi. In questo ambito, i metodi di relazione quantitativa struttura-attività (QSAR), l'analisi di *docking* molecolare e di dinamica molecolare, possono efficacemente contribuire ad una razionalizzazione dei risultati ottenuti per via sperimentale. Il principale obiettivo di tale approccio multidisciplinare è quello di determinare i dettagli conformazionali fondamentali per l'interazione tra la proteina ed il suo ligando, in modo da poter avere informazioni utili per la progettazione di nuovi potenziali farmaci. Le molecole individuate tramite tale studio necessitano sempre di ulteriori indagini sperimentali *in vitro* ed *in vivo* per avere un'esaustiva comprensione del meccanismo di azione del potenziale farmaco e per affrontare le problematiche che influenzano fortemente la farmacocinetica del composto o dei suoi metaboliti (solubilità, permeabilità delle molecole e possibili interazioni con altre biomolecole o con altri farmaci). Nell' ottica di definire il meccanismo di azione l'analisi a livello molecolare dei dettagli strutturali delle interazioni tra biomolecole permetterebbe l'acquisizione di informazioni utili per la loro comprensione e, perciò, per la determinazione delle funzionalizzazioni necessarie ad incrementare l'attività del potenziale farmaco. In questo ambito si inserisce questo lavoro di tesi, per il quale è stato utilizzato un approccio integrato e multidisciplinare di tecniche spettroscopiche (fluorescenza, dicroismo circolare (CD), risonanza magnetica nucleare (NMR) and *Dynamic Light Scattering*, (DLS)) e computazionali (*molecular docking*, *molecular dynamics*) per migliorare l'interazione specifica di molecole di interesse biotecnologico e terapeutico al fine di ottenere composti con attività antitumorale e proprietà farmacologiche potenziate rispetto a quelle con attività nota. A tal fine sono stati scelti due sistemi proteici per i quali esiste un'ampia letteratura e si dispone di banche dati relative a parametri strutturali e farmacologici. In particolare in questo lavoro sono state progettate e sintetizzate nuove molecole potenzialmente in grado di bloccare l'attività di due proteine sovraespresse in diverse patologie: Akt (trattata nella sezione ***Workpackage 1***), e CXCR4 (descritta nel ***Workpackage 2***) [2- 4].

Workpackage 1

Il **Workpackage 1** della mia tesi verte sulla progettazione e la sintesi di nuovi composti antiproliferativi. Tra questi si è focalizzata l'attenzione su quelle molecole in grado di bloccare l'attività di Akt, la cui sovraespressione è stata collegata ad una proliferazione non controllata delle cellule cancerogene. In tale contesto un recente lavoro evidenzia l'importanza dell'anello pirrolo [1,2-*a*]quinoxalinico come *scaffold* molecolare ed evidenziano il ruolo fondamentale di opportune sostituzioni in precise posizioni per modulare l'attività dell'intera molecola come agente antiproliferativo. Infatti, la presenza di un anello fenile ed un gruppo etil estere sulla porzione pirrolica di tale *scaffold*s si sono rivelate importanti per l'attività della molecola, come evidenziato dai saggi condotti su diverse linee cellulari tumorali umane [5]. Tali studi hanno inoltre mostrato la notevole influenza dei sostituenti benzilpiperidinilbenzimidazolone e fluorobenzimidazole sull' attività antiproliferativa.

In letteratura è riportato che, da un punto di vista sintetico, la formazione dell'anello phenyl-1*H*-pirrol-diestere costituisce la fase critica dell'intero processo. Per questo si è reso necessario confrontare i diversi protocolli sperimentali di sintesi chimica, già riportati per molecole analoghe, per valutare gli effetti di diverse condizioni (temperatura, solvente e catalizzatore) sulla resa e tempi di reazione. L'ottimizzazione delle condizioni di sintesi ha consentito di ottenere fenil-1H-pirrolo-diesteri, dai quali sono stati sintetizzati gli *scaffolds* pirrolo[1,2-*a*]chinossalinici diversamente sostituiti, che costituiscono gli intermedi chiave nella sintesi dei composti finali **1a-l**, le cui strutture sono schematizzate nella **Figura 1**.

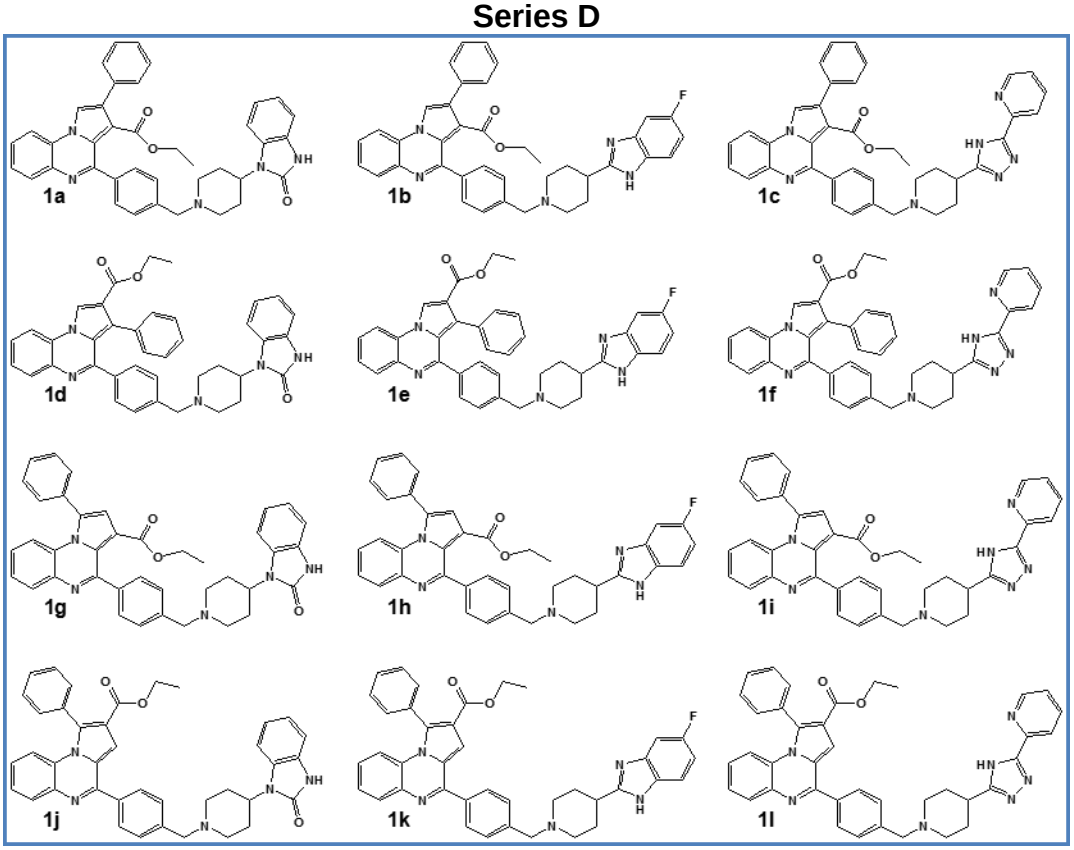


***Figure 1**. Struttura chimica dei composti **1a-l**.*

La struttura dei composti è stata confermata mediante caratterizzazione IR e NMR. Test di attività antiproliferativa (MTS) sono stati inoltre condotti su diverse linee cellulari cancerogene umane (U937, K562, Jurkat, U266 and HL60) al fine di confrontare diverse tipologie di cellule tumorali nelle quali la proteina Akt, normalmente coinvolta nella regolazione del processo di proliferazione, è sovraespressa. Le molecole **1a**, **1e**, **1g** e **1h** hanno mostrato un'elevata attività citotossica nei confronti di alcune linee cellulari tumorali (IC_{50} compresi nell'intervallo 3-9 μM) e una bassa tossicità contro cellule ematopoietiche normali (IC_{50}> 50 μM) (***Tabella 1***). Questi composti, mostrando un interessante indice di selettività, possono costituire candidati adatti per ulteriori studi farmacologici. Da un punto di vista generale della SAR, questi risultati biologici preliminari sottolineano l'importanza della presenza di un gruppo benzylpiperidinylfluorobenzimidazole nella posizione C-4 dello *scaffold* pirrolo[1,2-*a*] chinossalinico, ma anche la necessità di una funzionalizzazione sull' anello pirrolico. I dati biologici sembrerebbero indicare che il composto **1e** é il più promettente. Sarebbe, perciò, interessante valutare in maniera

approfondita l'attività biologica di tale molecola per poterne meglio definire il meccanismo di azione.

***Tabella 1**. Attività in vitro dei composti **1a-l** su linee cellulari K562, U937, HL60, Jurkate U266, e citotossicità su cellule mononucleari del sangue periferico umano PBMNC + PHA.*

IC_{50} values (μM)[a]						
Inibitore	**K562**	**U937**	**HL60**	**Jurkat**	**U266**	**Citotossicità su cellule mononucleari attive del sangue periferico umano (PBMNC) PBMNC + PHA**
A6730	17 ± 0.3	8 ± 0.2	5.5 ± 0.2	3.5 ± 0.2	n.d.[b]	>50
1a	4 ± 0.1	>50	>50	41 ± 1.2	>50	>50
1b	3 ± 0.1	21 ± 2.2	19 ± 2.8	5 ± 0.2	3.5 ± 0.1	14 ± 1
1c	14 ± 0.3	7 ± 0.3	12 ± 2.3	5 ± 0.1	>50	16 ± 2
1d	9 ± 0.3	>50	3.5 ± 0.1	4 ± 0.1	10 ± 0.8	41.3 ± 5
1e	3 ± 0.1	n.d.	10 ± 0.9	5 ± 0.2	18 ± 1.1	>50
1f	7 ± 0.2	7 ± 0.1	5 ± 0.1	2 ± 0.1	5 ± 0.1	8 ± 0.5
1g	8 ± 0.2	9 ± 0.4	6 ± 0.1	4 ± 0.1	9 ± 0.8	>50
1h	7 ± 0.3	4 ± 0.1	8 ± 0.2	6 ± 0.15	4 ± 0.1	50 ± 4
1i	12 ± 0.4	11 ± 0.9	24 ± 3	5 ± 0.1	8 ± 0.5	50 ± 6
1j	8.5 ± 0.3	>50	>50	6.5 ± 0.1	>50	>50
1k	3 ± 0.1	3 ± 0.1	3 ± 0.1	3.5 ± 0.1	3 ± 0.1	13 ± 1
1l	3.5 ± 0.1	8 ± 0.2	7 ± 0.3	3	3 ± 0.1	12 ± 1

[a]I valori di IC_{50}(μM) corrispondonoalla media ± deviazione standard rispetto a tre esperimenti indipendenti.
[b]n.d. = nondeterminato.

La possibilità di prevedere la tossicità delle molecole è molto importante nella progettazione di nuovi potenziali farmaci. In questo contesto, le molecole sintetizzate **1a-l** sono state analizzate mediante il *software* "**Molinspiration web-enabled**" per predire dei parametri chimico-fisici importanti per la farmacolgia di tali molecole, come la lipofilia (Clog*P*) e l'assorbimento intestinale (calcolo delle superfici polari molecolari, PSA). La superficie polare molecolare (PSA) dei composti della serie **1a-l** suggerirebbe il loro assorbimento intestinale (PSA < 140 $Å^2$). Inoltre tali molecole non sono in grado di attraversare la barriera emato-encefalica (PSA > 60 $Å^2$). Alla luce di queste predizioni, i composti **1a-l** presenterebbero una buona biodisponibilità per via orale con una potenziale bassa tossicità rispetto al sistema nervoso centrale. Successivamente, allo scopo razionalizzare i risultati biologici e determinare i dettagli strutturali che ne potrebbero essere responsabili nelle molecole, studi di *docking* sono stati condotti utilizzando il server Swiss Dock (http://www.swissdock.ch). La proteina considerata per l'analisi iniziale è stata la struttura cristallografica della proteina Akt intera depositata nel Protein Data Bank (cod. 3O96 [6]). I risultati di tale studio indicano la molecola **1e** come la migliore tra quelle sintetizzate, evidenziando che la presenza di un anello fenilico nella terza posizione e l'introduzione del piccolo

gruppo etil estere sulla seconda posizione dell'anello quinoxalinico potrebbero favorire una migliore interazione tra il ligando e la proteina, come mostrato in ***Figura 2***.

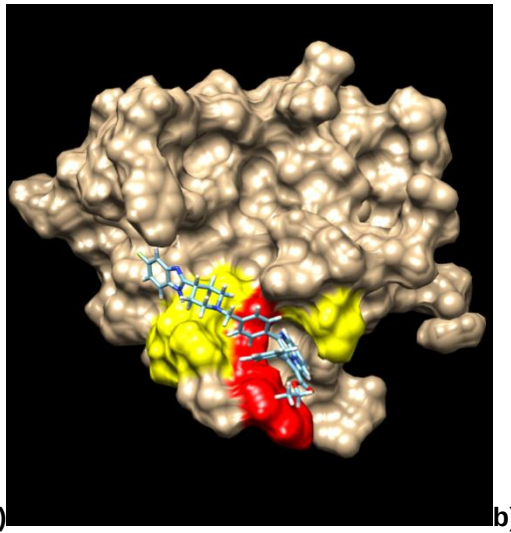
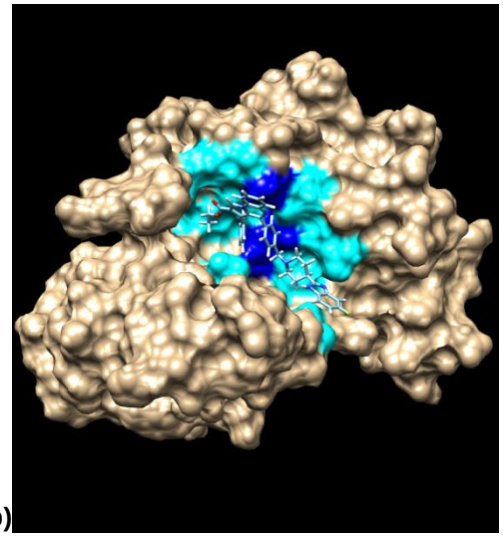
a)  b) 

Figura 2*. Migliore struttura ottenuta attraverso il docking del il composto* ***1e*** *rispetto alla Proteina intera Akt* ***(****cod. 3O96 [6]);* ***a)*** *Dominio PH: in rosso, i residui coinvolti in almeno 10 interazioni e in giallo, i residui coinvolti in al massimo 9 interazioni.* ***b)*** *Dominio chinasico: in blu i residui coinvolti in almeno 10 interazioni e in ciano, i residui coinvolti in al massimo 9 interazioni.*

Succesivamente, gli studi di *docking* sono stati condotti unicamante sul dominio PH estrapolato dalla struttura cristallografica di PKBα (codice PDB: 1UNR) [7]. Questa analisi è stata effettuata al fine di verificare se la nostra migliore molecola può essere considerata come un potenziale inibitore allosterico della proteina Akt o un semplice ligando PH-dipendente. I risultati di tale studio hanno mostrato un andamento simile a quello ottenuto per la proteina intera. In particolare, essi rivelano che **1e** è il miglior composto. Infatti, la sovrapposizione della sua migliore struttura ottenuta dal *docking* sul dominio PH di Akt e la struttura cristallina di inositol-(1,3,4,5)-tetrakisfosfato (IP4) in complesso con il dominio PH di Akt1 (codice PDB:. 1UNQ [7]) evidenzia la presenza del nostro ligando in una tasca distinta dal sito di legame dell'inositolo. In conclusione I risultati di *docking* indicano che le molecole **1a-l**, più che essere semplici ligandi PH dipendenti, costituiscono degli inibitori allosterici per via della loro abilità di interagire con entrambi i domini della proteina Akt, cioè il dominio PH e quello chinasico.

<u>Risultati del Workpackage 2</u>

Il **Workpackage 2** è focalizzato su CXCR4, un recettore accoppiato alla proteina G che è coinvolto nel *signaling* di differenti fattori. In un recente lavoro viene descritto il comportamento strutturale e l'attività biologica tra CXCR4 e il suo ligando naturale CXCL12 [8]. L' interazione CXCL12-CXCR4 svolge un ruolo fondamentale in malattie come HIV, cancro, sindrome WHIM, artrite reumatoide, fibrosi polmonare e il lupus, che è per questo indicata come un valido bersaglio terapeutico [9]. Considerato l'elevato potere e l'alta specificità che i peptidi possono avere [10], si registra, dai dati di letteratura recente, un interesse molto ben documentato nello sviluppo di strumenti terapeutici e diagnostici di natura peptidica. In particolare, i peptidi intrinsecamente disordinati (IDP) rappresentano attualmente un innovativo e interessante punto di partenza, perché il disordine conformazionale, inteso come mancanza di qualunque

elemento di struttura secondaria regolare, quale eliche e foglietti beta, consente al peptide di subire variazioni conformazionali solo in seguito all' interazione con i sistemi bersaglio. Tale capacità degli IDP conferisce loro una serie di vantaggi rispetto ai loro congeneri strutturati, quali maggiori superfici di interazione ed una migliore cinetica di legame. Infatti, tali peptidi esplorano dinamicamente un insieme di conformazioni assolutamente flessibili, adottando quelle più stabili solo dopo essersi legati ai ligandi [11]. Gli IDP possono, perciò, instaurare interazioni con elevata specificità / bassa affinità con i loro interattori. Alla luce di ciò, abbiamo disegnato due peptidi lineari, PeptideE (YGECPCE-allile) e PeptideK (YGECPCK- allile), con un elevato contenuto di residui in grado di favorire strutture disordinate [12], considerando come modello la regione N-terminale di CXCL12 e la sequenza CPC, dimostrata essere responsabile dell'interazione tra il peptide CXCL12 ed il recettore CXCR4[8]. A tali sequenze sono stati aggiunti, alle estremità C-terminali, amminoacidi carichi (Glu e Lys, rispettivamente per il PeptideE e PeptideK) allo scopo di analizzare l'effetto della carica sulle proprietà di legame del peptide. Il C-terminale di tali peptidi é stato inoltre modificato da un estere allilico per eventuali nuove funzionalizzazioni delle molecole. Gli spettri CD hanno evidenziato l'assenza di elementi di regolare struttura secondaria e quindi la natura disordinata di PeptideE e PeptideK. Tale risultato è stato confermato dalla bassa dispersione spettrale degli spettri 1D [1H] e dalla quasi totale assenza di segnale negli spettri 2D [1H, 1H] NOESY, il che indicherebbe per entrambi i peptidi una elevata flessibilità ed una natuta disordinata/non foldata. Dalla valutazione delle deviazioni dei *chemical shift* degli Hα protonici dai valori *random coil*, che risultano essere in prevalenza piccole e positive, è stata confermata la presenza di strutture estese disordinate. Infine, lo stato disordinato di PeptideE e PeptideK è stato confermato ulteriormente dai *pattern* ROE ottenuti, che mostrano contatti sequenziali (Hα*i*–HN*i+1*) forti, tipici di peptidi non organizzati strutturalmente. I risultati dell'indagine NMR confermano, perciò, la flessibilità di tali peptidi e, quindi, la loro impossibilità ad assumere conformazioni ben definite. Lo studio successivo di simulazioni di dinamica molecolare (MD) (***Figura 3***), condotto sui peptidi liberi, ha confermato tali risultati sulla flessibilità e l'assenza di una strutturazione, ma ha anche evidenziato la presenza di una rete di legami ad idrogeno intramolecolari temporanei e dinamici di tipo catena principale-catena principale (MM) e di interazioni con le molecole d'acqua.

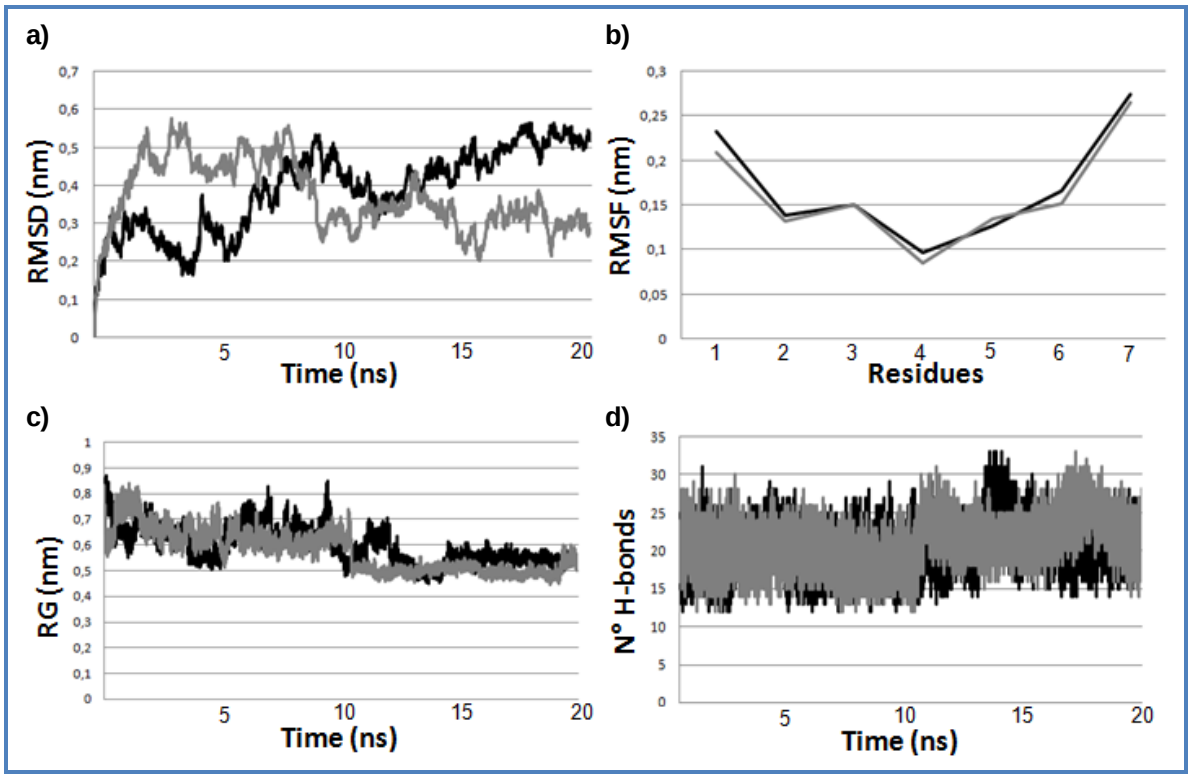


***Figure 3**. Analisi delle simulazioni di dinamica molecolare condotte su PeptideE (in nero) e su PeptideK (in grigio) a pH fisiologico in solvente esplicito: **a)** deviazione della radice quadratica media (RMSD), **b)** fluttuazioni della radice quadratica media (RMSF), **c)** raggio di rotazione (gyration radius), and **d)** numero di legami idrogeno.*

Pertanto, le strutture dei due peptidi possono essere rappresentate come insiemi conformazionali caratterizzati da strutture secondarie irregolari e fluttuanti. Infatti, i peptidi non assumono un'unica e definita conformazione, ma vanno continuamente da un *cluster* ad un altro (***Figura 4***). Pertanto essi non possono essere correttamente descritti da un'unica struttura ma da un "conformational ensamble".

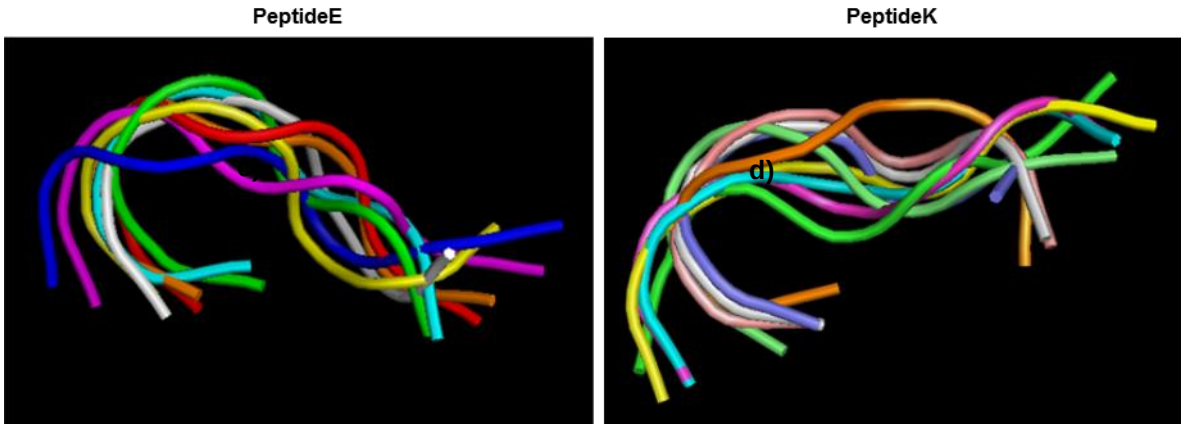


***Figura 4.** Sovrapposizione dei cluster più popolati ottenuti per il PeptideE (8 clusters) e la PeptideK (9 clusters) durante le loro dinamiche molecolare.*

Considerato che una delle maggiori difficoltà nella progettazione di un potenziale farmaco consiste nella sua veicolazione al sito di interesse, con tutte le problematiche ad esso connesse (come il superamento delle membrane cellulari), si è pensato di funzionalizzare l'estremità N-terminale dei peptidi E e K mediante code

idrofobiche e conferendo, pericò, una natura anfifilica ai peptidi liberi. I peptidi anfifilici sono largamente descritti in letteratura come ottimi canditati per il *drug delivery* di farmaci grazie alla la loro capacità di formare aggregati sopramolecolari (ad esempio vescicole e micelle), dovuta alla coesistenza di una porzione idrofilica ed una idrofobica. In tale contesto i peptidi anfifilici (PAs, ***Figura 5***) $(C18)_2$-PeptideE e $(C18)_2$-PeptideK sono stati ottenuti coniugando due catene alifatiche di 18 atomi di C all'estremità N-terminale di entrambe le sequenze. In aggiunta, sono stati sintetizzati altri due derivati PAs: $(C18)_2$-L-PeptideK e $(C18)_2$-L-PeptideE, caratterizzati dalla presenza di un *linker* etossilico (AhOh) tra la sequenza peptidica e la coda idrofobica (***Figura 5***). Questo *linker,* lo stesso per entrambi i peptidi, permette di aumentare l'idrofilicità dei peptidi senza modificarne la carica. In questo modo si limitano possibili interazioni elettrostatiche *intra* e *inter*molecolari. Inoltre, esso permette di allontanare il peptide bioattivo dal guscio idrofobico degli aggregati. La lunghezza relativamente piccola di questo spaziatore è stata scelta per evitare la formazione di una tasca idrofila, in grado di nascondere le sequenze amminoacidiche [13].

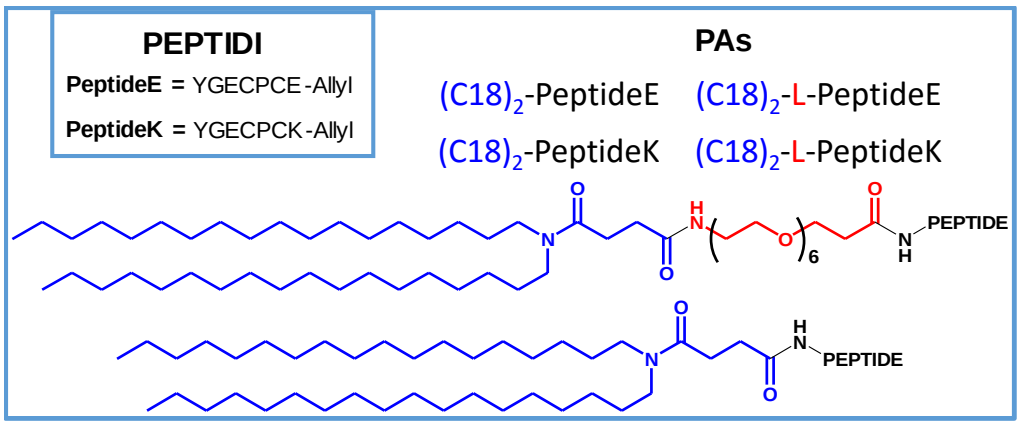


Figura 5. *PeptideE e PeptideK liberi e i loro derivati PAs. (In* ***blu*** *é riportata la coda idrofobica e in* ***rosso*** *il linker introdotti). La catena PEG è delimitata da parentesi.*

Gli aggregati sopramolecolari dei peptidi PAs sono stati ottenuti solubilizzando e disperdendo i derivati anfifilici del PeptideE in tampone fosfato 10 mM (pH 7.4) e quelli del PeptideK in tampone TRIS 10 mM (pH 8.0). La scelta del tampone da utilizzare è stata imposta dalla differente propensione ad aggregare dei peptidi PAs. Come dimostrato da studi di fluorescenza e DLS, tutti i derivati anfifilici progettati sono in grado di auto-assemblare in aggregati con un diametro medio compreso tra 80 e 110 nm, indipendentemente dalla carica all'estremità C-terminale e dalla presenza del linker. Le soluzioni di ogni PA aggregato mostrano una distribuzione monomodale, che indica la presenza di una sola popolazione di aggregati. In ***Tabella 2*** sono riportati i dati ottenuti dall' analisi di per fluorescenza e DLS di tutti i composti: concentrazione critica micellare (**CMC**), raggio idrodinamico (R_H), coefficienti di diffusione (***D***), e indice di polidispersità (**PDI**).

***Tabella 2**. Parametri strutturali ottenuti per gli aggregati da misure di fluorescenza (concentrazione critica micellare CMC) e misure DLS (raggio idrodinamico R_H, coefficienti di diffusione D, e indici polidispersità PDI).*

Peptide	CMC (mol kg^{-1})	R_H (nm)	D x 10-12 (m^2s^{-1})	PDI
$(C18)_2$-PeptideK	7.0 x 10^{-6}	80 ± 27	3.1 ± 1.0	0.270
$(C18)_2$-L-PeptideK	1.9 x 10^{-5}	82 ± 34	3.1 ± 1.5	0.182
$(C18)_2$-PeptideE	3.9 x 10^{-5}	112 ± 44	1.9 ± 0.7	0.264
$(C18)_2$-L-PeptideE	3.2 x 10^{-5}	110 ± 39	2.4 ± 0.5	0.226

Gli spettri CD ottenuti per i PAs derivati (acquisiti ad una concentrazione superiore a quella critica micellare) mostrano che la presenza del *linker* non influenza significativamente la struttura della loro porzione peptidica. Inoltre, tali spettri hanno evidenziato che la presenza di un residuo di Lys rispetto al residuo di Glu all'estremità C-terminale può indurre un riarrangiamento della struttura secondaria con un parziale *folding* della porzione peptidica, confermando l'importanza della carica in queste sequenze peptidiche. I composti da noi sintetizzati costituiscono i primi peptidi anfifilici contenenti la sequenza CPC in grado di aggregare e formare strutture sopramolecolari dotate di attività *in vitro.* Infatti, i test biologici, condotti sulle stesse cellule e alle medesime concentrazioni utilizzate per la molecola di riferimento AMD3100, hanno rivelato che, nonostante tutti peptidi non presentino una evidente affinità di legame per CXCR4 (dati non mostrati), essi sono in grado di indurre una produzione di cAMP più elevata di quella di AMD3100 e in maniera dose dipendente. Tali dati indicherebbero una significativa attività inibitoria dei peptidi sintetizzati verso CXCR4. I valori di cAMP di tutti i peptidi e quello relativo ad AMD3100 (molecola di riferimento) sono riassunti in ***Figura 6***, dove tutti i valori sono espressi come concentrazione percentuale rispetto a due concentrazioni (1μM e 10μM). Il peptide CXCL12, ligando naturale di CXCR4, è stato considerato in tale indagine come controllo.

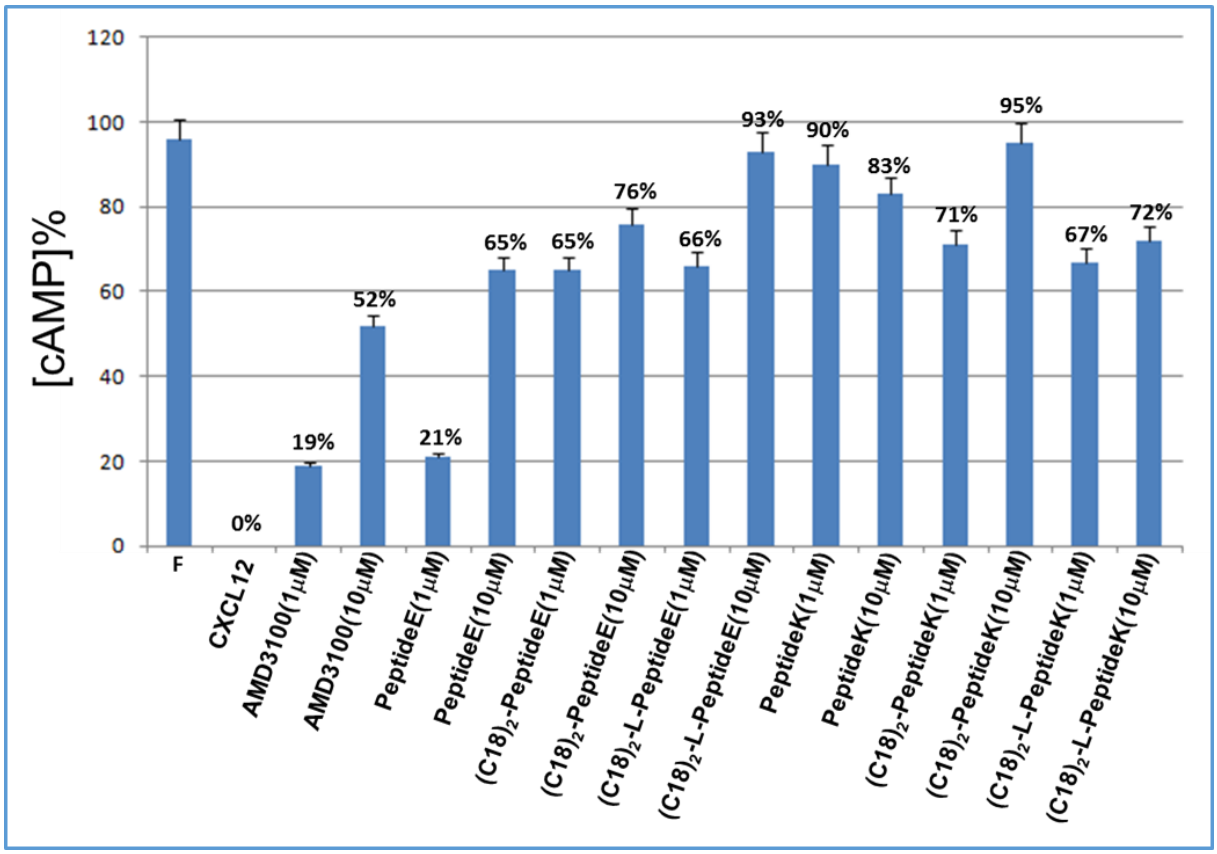


***Figure 6**. Confronto delle modulazioni di cAMP valutate per Fosfokolin (F), CXCL12, PeptideE, PeptideK, $(C18)_2$-PeptideE, $(C18)_2$-PeptideK, $(C18)_2$-L-PeptideE, $(C18)_2$-L-PeptideK e per la molecola di riferimento AMD3100, alle concentrazioni 1 μM e 10 μM.*

Sebbene questi risultati siano da considerare preliminari e consapevoli che sono necessarie ulteriori analisi, è possibile sicuramente affermare che questi risultati sono incoraggianti e mostrano la validità dell'approccio multidisciplinare alla base di questa indagine scientifica. Il passo successivo andrà nella direzione di apportare ulteriori modifiche chimiche, utilizzando la reattività del gruppo allil estere, per l'ottenimento di nuovi composti sopramolecolari di natura peptidica, in grado di essere incapsulati nelle cellule e, quindi, di veicolare il peptide bioattivo e/o sonde per terapie mediche personalizzate e/o per indagini di *imaging* a risonanza magnetica (MRI).

1. Paul S.M. *et al*., Nat.Rev. Drug Discov., 2010, 9(3), 203-214;
2. Yu Y. et al., Plos One, 2015, 10(10), e0140479;
3. Zabel B. A. *et al.*, J.Immunol., 2009, 183, 3204–3211;
4. Ma Q. *et al.,* Proc. Natl. Acad. Sci. USA, 1998, 95, 9448–9453;
5. Desplat V. et al., J.Enzym.Inhib. Med. Ch., 2011, 26(5), 657-667;
6. Wu-I W. et al., *Plos One*, 2010, 5(9), e12913;
7. Milburn CC et al, Biochem. J., 2003, 375, 531-538;
8. Costantini S. *et al.*, J.Pept. Sci., 2014, 20 (4), 270-278;
9. Tamamura H. e Fujii N., *Expert Opin. Ther. Tar.*, 2005, 9(6), 1267–1282;
10. Accardo A. *et al.*, Mol. BioSyst., 2013, 9, 1401-1410;
11. Mittal J., J. Phys. Chem. B, 2013, 117 (1), 118-124;
12. Uversky V. N., *Int.J.Biochem. Cell B.*, 2011, 43(8), 1090-1103;
13. Matsumura Y. & Maeda H., CancerRes., 1986, 46, 6387-6392.

Résumé

Ce travail de thèse PhD concerne l'application d'une approche intégrée caractérisée par l'utilisation de différentes techniques expérimentales et de calculs pour obtenir une meilleure compréhension des détails structurels en solution et des interactions intermoléculaires impliquées dans les mécanismes d'action de deux systèmes: Akt-ligand hétérocyclique (***Workpackage 1***) et la chimiokine, CXCR4 (***Workpackage 2***). Il est connu que la dysrégulation de Akt et CXCR4, protéines normalement chargées de la régulation des processus dans la cellule tels que la prolifération, est surexprimée dans différents cancers humains. Pour cette raison, ces protéines sont considérées des cibles importantes dans le développement d'agents thérapeutiques potentiels contre le cancer. Des efforts récents dans le développement et l'évaluation biologique de petites entités moléculaires inhibitrices d'Akt, une sérine/thréonine protéine-kinase, ont conduit à l'identification de nouveaux inhibiteurs hétérocycliques. Parmi les composés hétérocycliques qui ont attiré une certaine attention en raison de leur large activité biologique, le système hétérocyclique pyrrolo[1,2-*a*]quinoxaline a été identifié comme édifice moléculaire particulièrement intéressant pour l'activité antiprolifératives contre diverses lignées de cellules cancéreuses humaines. Ainsi, en tenant compte des résultats obtenus précédemment sur l'activité antiproliférative de dérivés de type pyrrolo[1,2-*a*]quinoxaline, une nouvelle série de ces dérivés **1a-l** a été conçue et préparée via une stratégie de synthèse hétérocyclique multi-étapes. Leur cytotoxicité a été évaluée sur cinq lignées cellulaires de leucémie, Jurkat et U266 (lignées lymphoïdes), puis K562, U937 et HL60 (lignées myéloïdes), ainsi que des cellules sanguines humaines mononucléaires normales. Cette nouvelle série a montré un potentiel cytotoxique prometteur contre toutes les lignées cellulaires leucémiques testées ; et certains composés ont montré une activité biologique supérieure à celle du composé de référence A6730. Parmi ces composés, les dérivés **1a**, **1e**, **1g** et **1h** semblent plus prometteurs en raison de leur haute activité contre ces leucémies et leur faible cytotoxicité contre des cellules hématopoïétiques normales. Des relations de structure-activité de ces nouveaux composés synthétiques **1a-l** sont également discutées. En outre, des résultats préliminaires menés sur Akt puis sur le domaine de PH isolé d'Akt, ont montré que ces nouveaux composés peuvent être considérés comme des inhibiteurs allostériques potentiels. La seconde partie des travaux concerne la conception et la synthèse de deux nouvelles séquences peptidiques contenant quelques acides aminés et présentant une propension au désordre. De plus, ils présentent une unité CPC: motif se liant au ligand spécifique de CXCR4 à savoir la liaison CXCL12. Les structures préférentielles de ces peptides ont été analysées par les techniques CD et RMN qui ont fait ressortir leur flexibilité. Ces résultats ont également été confirmés par des simulations de MD qui ont démontré la capacité de ces peptides à assumer des ensembles de conformations stabilisées par un réseau de liaisons hydrogène, et caractérisées par des interactions avec des molécules d'eau. Ensuite, nous avons étudié la possibilité que ces peptides puissent être utilisés comme têtes polaires reliées à alkyl chaînes pour générer de nouveaux édifices moléculaires (peptides amphiphiles nommés, PAs) pour la vectorisation de médicaments dans les cellules surexprimant le récepteur CXCR4. Pour cela, nous avons étudié dans un premier temps l'effet de la chaîne alkyle sur le réarrangement de conformation des peptides intrinsèquement désordonnés. Des études de fluorescence et de DLS ont été réalisées pour évaluer les concentrations micellaires critiques et la dimension des agrégats supramoléculaires. Les tests biologiques ont souligné qu'une partie hydrophobe reliée à nos peptides pourrait améliorer leur activité sur CXR4. Ces édifices moléculaires montrent ainsi des activités prometteuses, voire plus que la molécule de référence (AMD3100). En nous basant sur ces résultats préliminaires, ces systèmes de peptides amphiphiles peuvent être en outre modifiés pour réaliser des composés supramoléculaires à base de peptides, susceptibles d'être encapsulés par des cellules vivantes pour une application biomédicale.

Introduction

It is well known that the development of new drugs requires different steps, such as synthesis of new molecules and screening of their activities to elaborate new libraries of potential therapeutic agents. In this context, the modern computational techniques may contribute to this complex process. This last is very challenging from the point of view of the associated cost and time, 1-2 billion of dollars and 12-13 years, respectively [1]. In this context, QSAR methods, docking analysis and MD tecniques, can effectively contribute to a rationalization of the results obtained experimentally.

The analysis at molecular level of conformational details involved in the interactions between biomolecules may give informations is of great interest in many biotechnological fields. In fact, it provides useful information for the design of molecules able to mimic the binding sites and thus, presenting potential therapeutic applications. Within this context, my PhD thesis has been focused on the application of an integrated approach composed by different experimental and computational techniques, such as molecular docking, MD (Molecular Dynamics), NMR (Nuclear Magnetic resonance), DLS (Dynamic Light Scattering) to have a better comprehension about the solution structure and the intermolecular interactions of two remarkable systems: Akt-heterocyclic ligand and CXCR4-chemokine. Both systems involve molecules that play important roles in physiological and pathological processes. The main goal of our investigation in solution was the determination of the protein conformational features that could be considered as fundamental for the interaction with the corresponding ligand. It is well known that protein-peptide and protein-small molecule interactions are mediated by specific domains. The analysis of the binding domain surface is of great interest in many biotechnological fields, as it provides useful information for the design of new molecules able to mimic the natural binder, thus presenting the potential application in the development of therapeutic agents. For this reason, we are focused in the design and synthesis of new molecules direct in the action of Akt (described in detail in the **Workpackage 1**) and CXCR4 (explained in the **Workpackage 2**), whose dysregulation is correlated to different pathologies. The activity of these molecules in the cells was evaluated by carrying out different types of biological tests. The analysis at molecular level of the interactions between these molecules and their natural target was conducted by molecular docking: in order to find a correlation between the ligand structure and its activity. The two considered systems and the respective protocols of analysis are reported below:

Workpackage 1: Akt-heterocyclic ligand

1. Chemical Synthesis
2. NMR, IR, MS, X-ray crystallography (compound structural characterization)
3. MTS cell proliferation assay (cytotoxicity test)
4. Calculations of Clog P, Topological Polar Surface Area (TPSA), number of hydrogen bond acceptor (nON) and donor (nNH/OH) atoms (computational prediction of toxicity and drug relevant properties)
5. Docking studies (structural details important in the Protein-ligand interaction)

Workpackage 2: CXCR4-Intrinsically Disordered Peptide ligand

1. Chemical Synthesis
2. Fluorescence and DLS (aggregates formation prediction and their characterization)

3. NMR and Circular dichroism (conformational studies of peptide)
4. Molecular modeling and dynamic simulations (Intrinsically disorder propensity of the peptides)
5. Migration Assay, Binding Assay and cAMP Assay (functional characterization of the peptides on CXCR4 receptor)

1. Workpackage 1: Akt-heterocyclic ligand system

1.1 Background

1.1.1. Akt

Akt is a member of the serine/threonine AGC protein kinase family involved in cellular metabolism, growth, proliferation and survival. To date, three mammalian isoforms (Akt1/PKBα, Akt2/PKBβ, and Akt3/PKBγ) of the kinase have been identified that share a high degree of sequence and structural homology, even though a possible distinct role has been suggested for each of them [2-4]. (***Figure 1***)

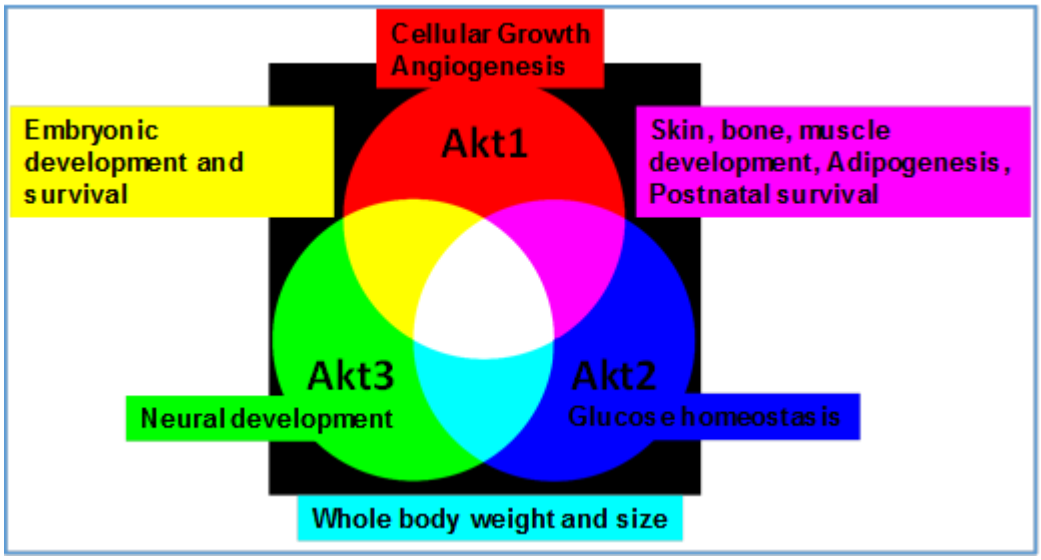


Figure 1. *Overlapping and specific functions of the Akt isoforms.*

They have a N-terminal pleckstrin homology domain (PH) mediating protein-protein and protein-lipid interactions, inter-domain linker, kinase domain and a 21-residue carboxy C-terminal hydrophobic motif (HM) [5]. The basic activation process of the Akt isoforms appears to be uniform and is mediated by growth factors [6]. (***Figure 2***)

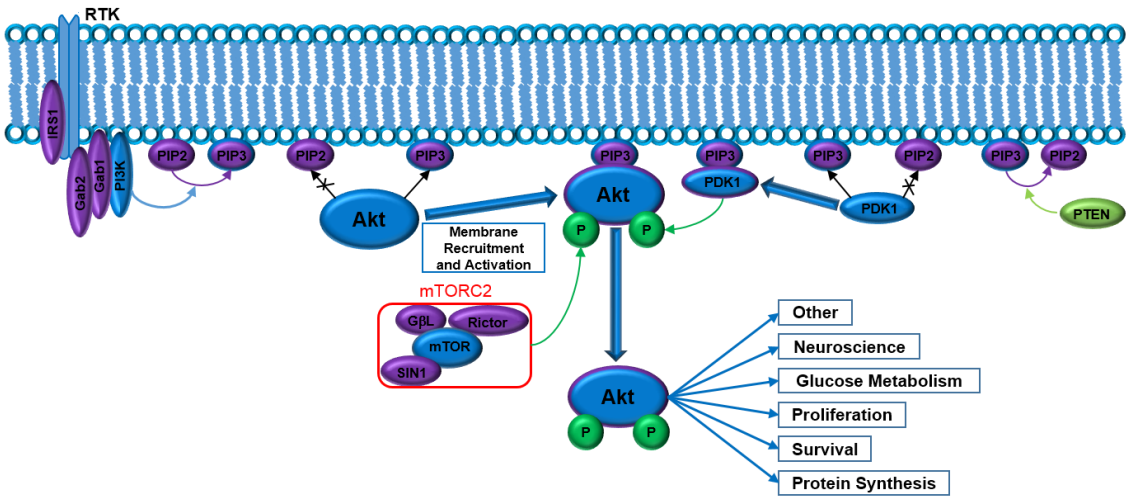


Figure 2*. Akt activation mechanism. PI3K protein convert PIP2 to PIP3 on the membrane surface. Akt and PDK1 protein translocate move to the membrane by PIP3. mTORC2 protein complex and PDK1 phosphorylate the Ser 473/474/472 and the Thr 308/309/305, respectively.*

It requires translocation from the cytosol to the plasma membrane followed by phosphorylation of both the threonine 308/309/305 and serine 473/474/472 residues (Akt1/Akt2/Akt3 respectively) which are regulated by multiple components of the well-known PI3K/AKT/mTOR pathway. In this pathway, PI3K phosphorylates phosphatidylinositol-(3,4)-bisphosphate (PIP2) to phosphatidylinositol-(3,4,5)-triphosphate(PIP3). The pleckstrin homology (PH) domains of both PI3K-dependent kinase-1 (PDK1) and AKT bind to PIP3 (located at the plasma membrane) thus allowing PDK1 to phosphorylate the activation loop of AKT at Thr 308/309/305. The activation process is completed with the phosphorylation of AKT at the C-terminal Ser 473/474/472 of the hydrophobic motif which occurs via the mTORC2/rictor complex. Once activated, AKT phosphorylates downstream target proteins including PRAS40, GSK3 (glycogen synthesis kinase 3), FOXO, BAD, endothelial nitric-oxid synthase and forkhead transcription factor (FHKR) [6]. (***Figure 3***)

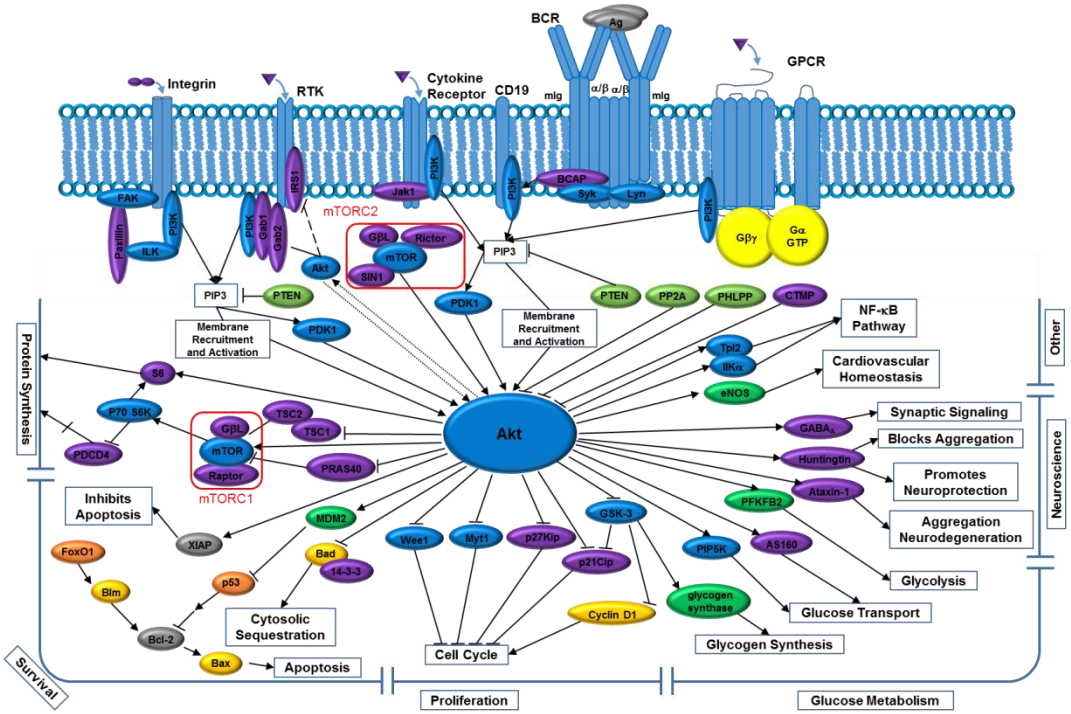


Figure 3. *Schematic representation of the complex Protein network in which Akt play a crucial role by which performs its different functions.*

Additionally, AKT also suppresses the inhibitory effect via phosphorylation of TSC1 and TSC2, thereby activating mTORC1 and ultimately leading to phosphorylation of 4E-BP1 and S6K. Dysregulation of PI3K/AKT/mTOR signaling pathway regulates many normal cellular processes, including cell proliferation, survival, growth and motility, processes that are critical for tumorogenesis. Indeed, the role of this pathway in oncogenesis has been extensively investigated and altered expression or mutations of many components of this pathway have been implicated in human cancer. Thus, the kinase Akt represents an attractive target for drug development [7].

1.1.2. Akt inhibitors

Different types of Akt inhibitors are reported in literature, which differ in their mode of action. There are inhibitors that interact with the PH domain of Akt and prevent the anchoring of the protein to the plasma membrane and, therefore, its activation such as Akt-in, a peptide that encompasses the βA strand of the proto-oncogene TCL1, Akt kinase coactivator. [8; 9] There are other ones, ATP-competitive inhibitors, which prevent the action of the already activated Akt [10]. They are described as non-selective against AKT isozymes, and poorly selective against closely related kinases. Efforts to identify new AKT specific and isozyme-selective inhibitors resulted in the discovery of novel selective allosteric AKT inhibitors [11; 12].

1.1.3. Akt allosteric inhibitors

The allosteric inhibitors bind non-covalently the enzyme in a specific site, different from the active one. This interaction induces a reversible conformational change which makes the active site no more able to interact with the substrate. Recent studies have highlighted the presence of an equilibrium between an inactive Akt conformation with the PH domain folded on the kinase one (named “PH-in”), and an active form with the PH domain moved from the kinase one (named “PH-out”). The allosteric AKT inhibitors prevent both membrane association and activation by phosphorylation. [5] Different examples of Akt allosteric inhibitors are already reported. Among them, A6730 is one of the most interesting compound (***Figure 4***). A crystal structure of the complex between this inhibitor and the Akt1 protein is deposited on the Protein Data Bank with the code 3O96. [5] A recent work shows the conformation of the ATP binding site in the AKT1:Inhibitor A6730 crystal structure (PDB code: 3O96 [5]). This study clearly reveals that the closed “PH-in” conformer is unable to bind ATP or ATP-competitive inhibitors. Not only does Phe 293 block the site, but critical ATP binding site residues interact with PH domain residues. Therefore, it was hypothesized in the same study that ATP-competitive AKT inhibitors are unable to bind the ‘PH-in’conformer and only bind the membrane-associated and phosphorylated ‘PH-out’ form, which has a properly configured ATP binding site. When the competitive inhibitor occupies the ATP binding site, the PH domain can not fully close onto the kinase domain and the phospholipid binding site remains exposed thus enhancing the propensity of AKT to localize to the membrane. In contrast, the AKT1:Inhibitor A6730 co-crystal structure shows the allosteric inhibitor locking AKT into a closed conformation with its phospholipid binding site blocked by thekinase domain. As a result, allosterically inhibited AKT remains cytosolic and is not activated via phosphorylation. In conclusion in this work, it was proved that ATP competitive inhibitors bind to AKT only when the protein assumes the open conformation. For this reason, there is a strong interest in the design of allosteric inhibitors that favor the inactive conformation [5].

1.1.4. Pyrrolo[1,2-*a*]quinoxaline ring as potential scaffold

In this context heterocyclic compounds are described as fundamental for the design and synthesis of potential Akt inhibitors. Among them, the pyrrolo[1,2-*a*]quinoxaline heterocyclic framework could constitute an useful basis. We have recently published three series (Series **A-C**) of new interesting substituted pyrrolo[1,2-*a*]quinoxalines (***Figure 4***) endowed with good activity towards the human leukemia cells [13-15]. These antiproliferative pyrrolo[1,2-*a*]quinoxaline derivatives have been previously designed as novel structural analogues of compound **A6730**, a well-described Akt inhibitor that presents antiproliferative activity against different human leukemia cell

lines [13-16]. Continuing our efforts in this field and considering the pharmacological activities of pyrroloquinoxalines on human leukemic cells, a new series (Series **D**) was designed and synthesized. Thus, by taking into accounts the best results obtained in series **B** (***Figure 4***), we decided to use the JG576 and JG572 pyrrolo[1,2-a]quinoxaline moieties as a template for the design of new derivatives **1a-l** in which the pyrrole nucleus is substituted in different positions by a phenyl and an ester function (**Series D**, ***Figures 4-5***). In relation to a previous works, further pharmacomodulations on the piperidine core have been considered, such as the introduction of new substituted heterocyclic systems [13-15]. The antiproliferative profile of the obtained derivatives **1a-l** was then evaluated *in vitro* against a panel of myeloid or lymphoid leukemic cell lines: U937, K562, Jurkat, U266 and HL60. Moreover, to determine their respective cytotoxicity, the new ethyl 4-[4-(4-substitutedpiperidin-1-yl)]benzyl-phenylpyrrolo[1,2-a]quinoxaline-carboxylate derivatives **1a-l** were tested on activated human peripheral blood mononuclear cells. Structure-activity relationships of these new synthetic compounds **1a-l** are here discussed. Finally, we used simple computational programs to predict the drug-like characteristics through the calculated physicochemical and toxicological properties of these new ethyl 4-[4-(4-substitutedpiperidin-1-yl)]benzyl-phenylpyrrolo[1,2-a]quinoxaline-carboxylate derivatives to determine their potential as new anti-leukemia drugs.

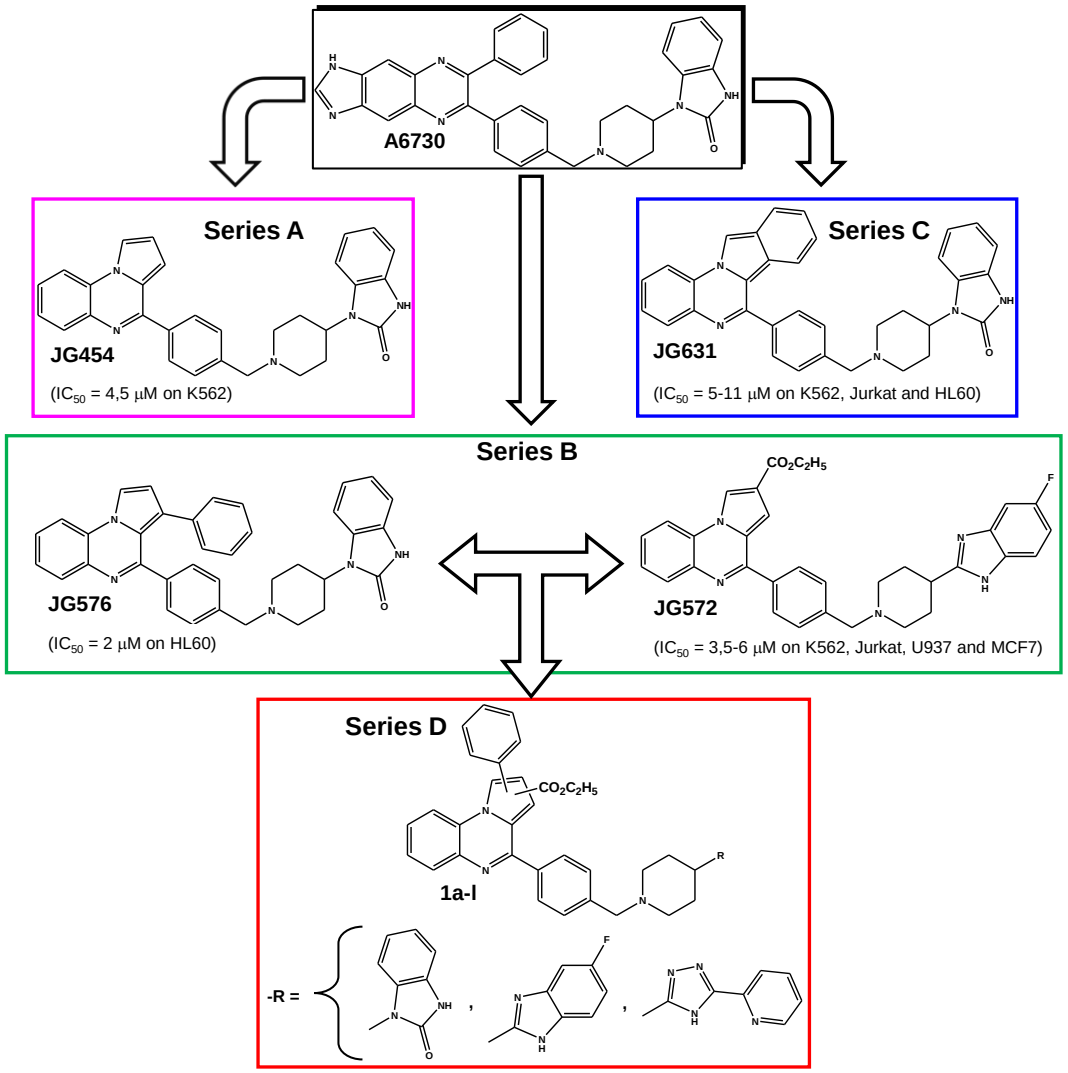


Figure 4. *Structure of bioactive compounds of previously described series A-C, and general structure of new synthesized substituted pyrrolo[1,2-a]quinoxaline derivatives* ***1a-l*** *(series D).*

1.2. Aim of Workpackage 1

The **Workpackage 1**of my thesis relies on the design and synthesis of new potential antiproliferative compounds. A recent work shows the importance of the pyrrolo[1,2-*a*]quinoxaline ring as scaffold and the relevance of the introduction of precise functionalizations to modulate its activity. In this context the presence of a phenyl ring and an ethyl ester group on the pyrrol moiety of the pyrrolo[1,2-*a*]quinoxaline ring seem to be important as revealed by the biological test conducted on different human cancer cell lines [13-15]. These results also showed the significant influence of a benzylpiperidinylbenzimidazolone and fluorobenzimidazole substituents for the activity of the compounds. As demonstrated in a previous work the critical step of the entire synthesis is the formation of the phenyl-1*H*-pyrrole-diesters ring. For this reason, different protocols reported in literature were compared in order to study the effect of different conditions (temperature, solvent and catalyst) on the reaction yield. The purpose was to determine the critical elements of the protocolsable to modulate significantly the formationof the phenyl-1*H*-pyrrole-diesters ring. Afterwards this structure was considered as starting material to synthesize the pyrrolo[1,2-*a*]quinoxaline scaffolds differently substituted **6a-d,** key intermediates in the synthesis of the final compounds **1**.Different methodologies were used to obtain the different molecules. In this context, a classical multi steps strategy was compared to a methodology characterized by a one-pot reduction-cyclization step. The goal was to verify the effectiveness of the presence of an entropically favoured step on the entire process of formation of the molecules **6a-d**. After this the final compounds **1a-l** were obtained by using different strategies to introduce a long chain in the position 4 of the pyrrolo[1,2-*a*]quinoxaline scaffold. In different works the presence of a spacer between the scaffold and the long chain is considered useful to improve the activity of the molecules. For this reason, a benzyl piperidine spacer was introduced to separate the two portions of the molecule able to modulate their potential antiproliferative activity. The benzylpiperidinylbenzimidazolone and thebenzylpiperidinyl fluorobenzimidazole are described in literature as possible chain able to influence the antiproliferative activity of this scaffold. These modifications were compared to another one characterized by the presence of pyridine ring bound to a triazole structure in a precise way. The aim was the determination of a correlation between the rigidity of the aromatic group and the activity of the molecules. As reported in many works the microwaves could be able to transfer energy to the reaction mixture in a better way if compared to the conventional way of heating. Considering this, we evaluated the possibility to use this mode of heating to verify the possibility to improve the time and the yield of the reaction. NMR experiments were conducted in order to confirm the structure of the synthesized molecules. The MTS cell proliferation assay (Promega, France) is a colorimetric assay system, which measures the reduction of a tetrazolium component (MTS) into formazan produced by the mitochondria of viable cells. This assay was used in order to evaluate the antiproliferative activity of our new synthesized compounds **1a-l**. Different human cancer cell lines were considered for the biological tests in order to compare different typologies of cancer cells where the Akt protein, normally involved in the proliferation regulation process, is overexpressed. The prediction of the possible toxicity of the molecules is very important in the design of new potential drugs. Within this context, the Molinspiration web-enabled software was used to predict their drug-like characteristics, such as the lipophilicity (calculating the Clog *P*) and the intestinal absorption (calculating the molecular polar surface areas, PSA). Afterwards, docking studies were conducted by using the Swiss Dock server

(http://www.swissdock.ch) [17; 18]. The protein input considered for the analysis was the crystal structure of the Akt entire protein. (pdb code: 3O96 [5]). The aim was the determination of structural elements in the chosen scaffold, in the spacer and in the introduced functionalizations important for the interaction between the ligand and the protein. Next, docking studies were carried out by considering only the Akt PH domain and the best scored molecules. The input target considered was the PH domain extracted from the crystal structure of PKBα (pdb code: 1UNR [19]). This analysis was made in order to verify if our molecule can be considered as a potential allosteric inhibitor of the Akt protein or a simple PH domain binder. All informations acquired by docking studies were considered to set a possible correlation between the structure of the molecules and their biological activities.

1.3. Results and discussion

1.3.1. Synthesis

Series D

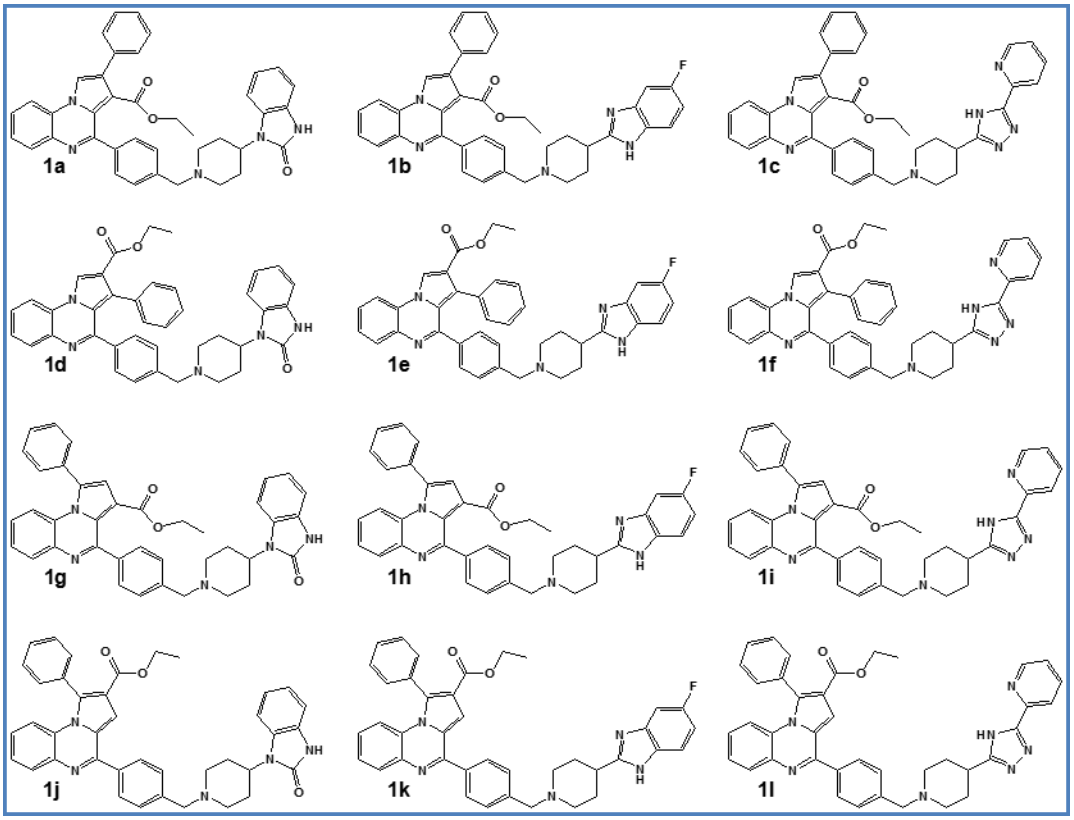


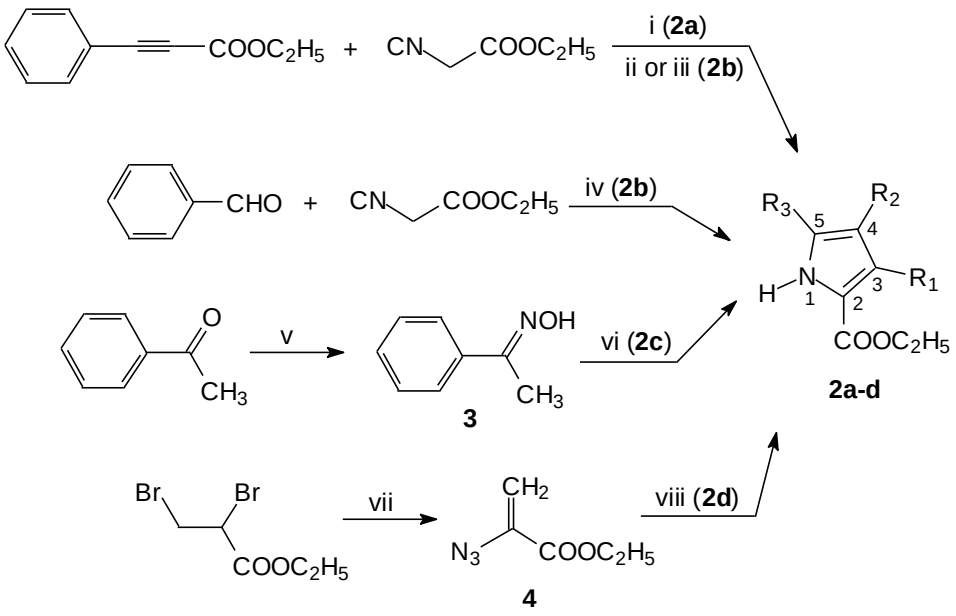
Figure 5. *Chemical structure of compounds* ***1a-l*** *obtained by different functionalizations.*

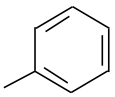
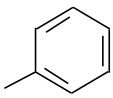
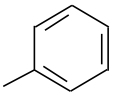
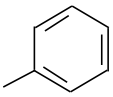
All reported pyrrolo[1,2-*a*]quinoxaline derivatives **1a-l** were synthesized from various substituted phenyl-1*H*-pyrrole-dicarboxylic acid ethyl ester **2a-d** (**Schemes 1-2**). Different strategies using classical or microwaves heating were considered for the synthesis of the phenyl-1*H*-pyrrole-diesters **2a-d** in order to introduce the phenyl and ester functions on the pyrrole ring (**Table 1**). The synthesis of the diethyl 4-phenyl-1*H*-pyrrole-2,3-dicarboxylate **2a** has been accomplished by treatment of ethyl isocyanide on ethyl phenylpropiolate under 1,3-bis(diphenylphosphino)propane (dppp) catalysis via a formal [3+2] cyclo-addition [20; 21]. (***Scheme 1***, ***Table 1***) Various attempts were investigated for the preparation of the diethyl 3-phenyl-1*H*-pyrrole-2,4-dicarboxylate **2b**. At first, this pyrrole **2b** was prepared by silver-catalyzed cycloaddition of commercialy available ethyl phenylpropiolate with ethyl isocyanoacetate in 1,4-dioxane or DMF at 80°C (***Scheme 1***, Methods A-C, ***Table 1***) [22; 23]. The very low yields obtained (7-12%, ***Table 1***) using classical or microwaves heating led us to investigate other methodologies. The copper-catalyzed reaction of ethyl isocyanide with the electron-deficient alkyne, ethyl phenylpropiolate, gave the diethyl 3-phenyl-1*H*-pyrrole-2,4-dicarboxylate **2b** with 58 to 62% in dioxane at 100°C using classical or microwave heating (**Table 1**) [24; 25]. Pyrrole **2b** could be also synthesized by reaction of ethyl isocyanoacetate with benzaldehyde in the presence of 1,8-diazabicyclo[5.4.0]undec-7-ene (DBU) in THF [26; 27].

Table 1. *Comparison of microwave-assisted synthesis of phenyl-1H-pyrrole-diester* ***2a-d*** *under standard reaction conditions.*

	Reagents and conditions	***Conventional heating***		***Microwaves heating (200W)***	
		Time	Yield (%)	Time	Yield (%)
2a	(i) dppp, dioxane, 100 °C	2 h	87	2 h	88
2b	(ii) *Method A*: Ag_2CO_3, dioxane, 80 °C	1 h	9	—	—
	(ii) *Method B*: 1) Ag_2CO_3, dioxane, 25 °C ; 2) 80 °C	1) 5 min 2) 30 min	10	1) 5 min 2) 30 min	7
	(ii) *Method C:* 1) Ag_2CO_3, NMP, 25 °C ; 2) 80 °C	—	—	1) 5 min 2) 30 min	12
	(iii) *Method D :* Cu_2O, phenanthroline, dioxane, 100 °C	2 h	58	40 min	62
	(iv) *Method E :* DBU, THF, 50 °C	5 h	21	1 h	35
2c	(vi) 1) $H_5C_2OOC\text{-}C{\equiv}C\text{-}COOC_2H_5$, DABCO, toluene, 80 °C; 2) 170 °C, P	1) 6 min 2) 2 h	28	1) 6 min 2) 45 min	28
2d	(viii) $C_6H_5\text{-}CO\text{-}CH_2\text{-}COOC_2H_5$, $Mn(OAc)_3{\cdot}2H_2O$, 40 °C	2 h	42	—	—

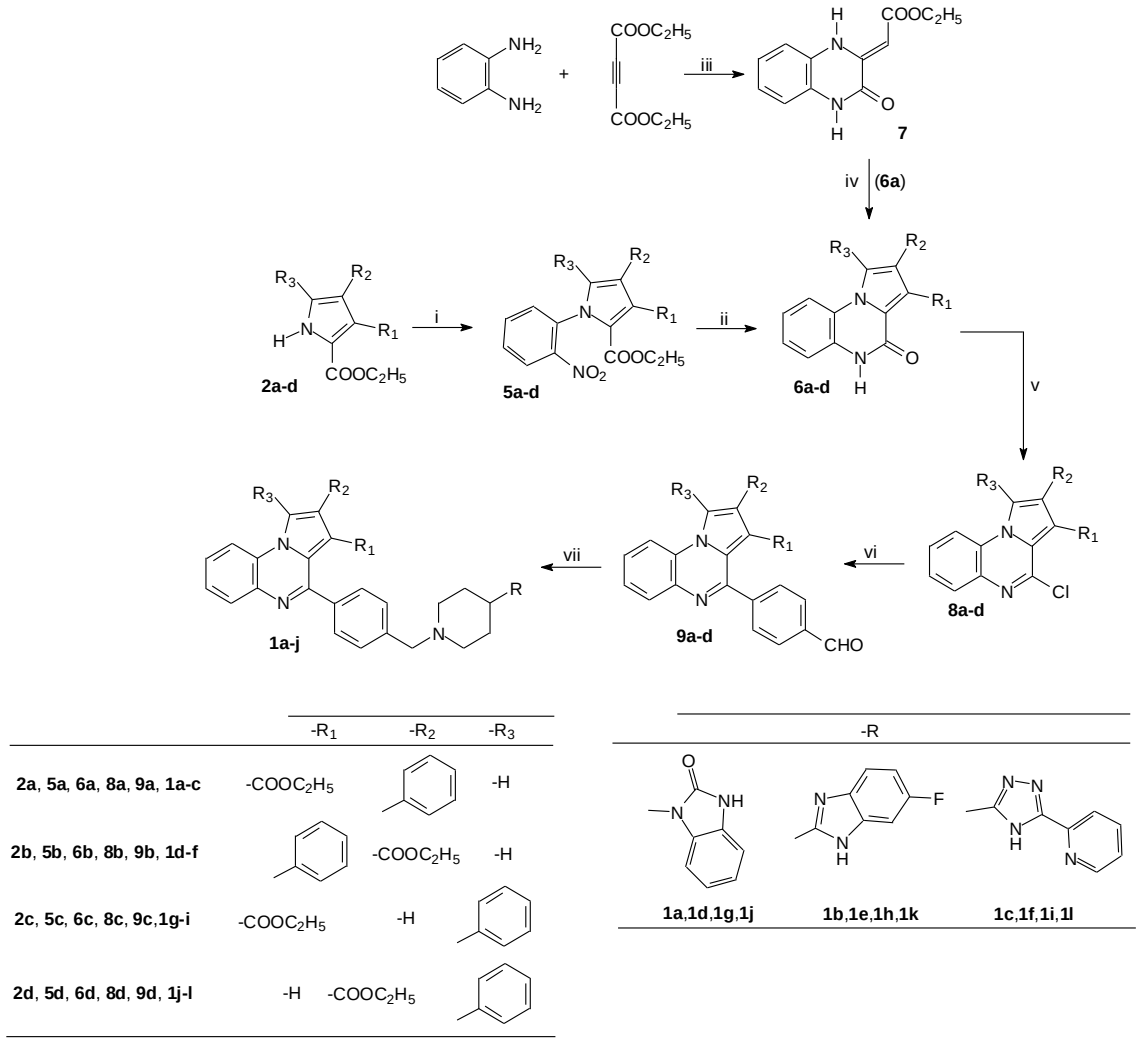
The synthesis of the diethyl 5-phenyl-1*H*-pyrrole-2,3-dicarboxylate **2c** has been accomplished in two steps starting from commercially available acetophenone via its oxime **3** (***Scheme 1***). The acetophenone oxime **3**, synthesized by the reaction of hydroxylamine on acetophenone [28], was then reacted with diethyl acetylene in the presence of 1,4-diazabicyclo[2.2.2]octane (DABCO) to form pyrrole **2c** via a thermal rearrangement by Trofimov reaction [29; 30]. Diethyl 5-phenyl-1*H*-pyrrole-2,4-dicarboxylate **2d** was prepared by manganese(III)-catalyzed formal [3+2] annulation of ethyl 2-azidoacrylate **4** and ethyl 3-oxo-3-phenylpropanoate [31]. The 2-azidoacrylate **4** was previously synthesized by treatment of ethyl 2,3-dibromopropanoate with three equivalents of sodium azide (NaN_3) in aprotic polar solvent such as dimethylformamide (DMF) [32; 33].



	-R_1	-R_2	-R_3
2a	-$COOC_2H_5$		-H
2b		-$COOC_2H_5$	-H
2c	-$COOC_2H_5$	-H	
2d	-H	-$COOC_2H_5$	

Scheme 1. *Synthesis of phenyl-1H-pyrrole-diesters **2a-d**; Reagents and conditions : (i) dppp, dioxane, 100 °C; (ii) Method A: Ag_2CO_3, dioxane, 80 °C; Method B: 1) Ag_2CO_3, dioxane, 25 °C ; 2) 80 °C; Method C: 1) Ag_2CO_3, NMP, 25 °C ; 2) 80 °C; (iii) Method D : Cu_2O, phénanthroline, dioxane, 100 °C; (iv) Method E : DBU, THF, 50 °C; (v) H_2NOH, HCl, Pyridine, EtOH, reflux; (vi) 1) H_5C_2OOC-C≡C-$COOC_2H_5$, DABCO, toluene, 80 °C; 2) 170 °C, P; (vii) NaN_3, DMF, 65 °C; (viii) C_6H_5-CO-CH_2-$COOC_2H_5$, $Mn(OAc)_3{\cdot}2H_2O$, 40 °C.*

The preparation of *N*-aryl pyrroles **5a-d** were obtained by nucleophilic substitution of the various pyrrole-2-carboxylates **2a-d** with 2-fluoro-nitrobenzene using cesium carbonate as the base in refluxing DMF solution (***Scheme 2***) [14; 15]. The preparation of **5a-d** was also performed under microwave irradiation. Reduction of the nitro moiety with iron in hot glacial acetic acid produced the spontaneous ring closure onto the ester in position 2 of the pyrrole moiety to afford the desired tricyclic pyrrolo[1,2-*a*]quinoxalines **6a-d** through a one-pot reduction-cyclization step [14; 15].



Scheme 2. *Synthesis of pyrrolo[1,2-a]quinoxalines* ***1a-l****; Reagents and conditions : (i) 2-fluoro-nitrobenzene, Cs_2CO_3, DMF, Δ; (ii) Fe, CH_3COOH, Δ; (iii) CH_3CN, RT; (iv) β–nitrostyrene, p-TsOH, Δ (v) $POCl_3$, Δ; (vi) OHC-C_6H_4-B(OH)$_2$, Pd[P(C_6H_5)$_3$]$_4$, K_2CO_3, toluene, EtOH, Δ; (vii) 4-(2-ketobenzimidazolin-1-yl)piperidine or 4-(5-fluorobenzimidazolin-2-yl)piperidine or 2-(3-piperidin-4-yl-1H-1,2,4-triazol-5-yl]pyridine, $NaBH_3CN$, MeOH, Δ.*

Sanaeishoar *et al.* previously reported the synthesis of lactam **6a** involving a three-component reaction between *o*-phenylenediamine, acetylenic ester and *β*-nitrostyrene using *p*-toluene sulfonic acid (PTSA) as a catalyst [34]. By following their procedure, the reaction only led to the formation of the intermediate ethyl 2-[3,4-dihydro-3-oxoquinoxaline-2(1*H*)-ylidene] acetate **7** [35; 36]. To convert this intermediate **7** into the lactam **6a**, we decided to modify the reported experimental procedure by heating quinoxaline **7** with *β*-nitrostyrene activated by PTSA in refluxing acetonitrile. This modified methodology afforded lactam **6a** with a 51% yield. The lactames **6a-d** were subsequently chlorodehydroxylated with phosphorous oxychloride, leading to the 4-chloroquinoxalines **8a-d**. 4-(Pyrrolo[1,2-*a*]quinoxalin-4-yl)benzaldehydes **9a-d** were easily prepared by a direct Suzuki-Miyaura cross-coupling reaction of 4-chloropyrroloquinoxalines **8a-d** with 4-formylphenylboronic acid performed in the presence of $Pd(PPh_3)_4$ as a catalyst, and in the presence of potassium carbonate used as the base [13-15]. The aldehydes **9a-d** were then engaged in a reductive amination with $NaBH_3CN$ and 4-(2-ketobenzimidazolin-1-

yl)piperidine or 4-(5-chloro-2-ketobenzimidazolin-1-yl)piperidine or 2-(3-piperidin-4-yl-1*H*-1,2,4-triazol-5-yl)pyridine to give the pyrrolo[1,2-*a*]quinoxalines **1a-l** [13-15].

1.3.2. X-ray structures

X-ray single crystal analysis was also performed on phenyl-1*H*-pyrrole-diesters **2b-d** in order to confirm their structures (***Figure 6***).

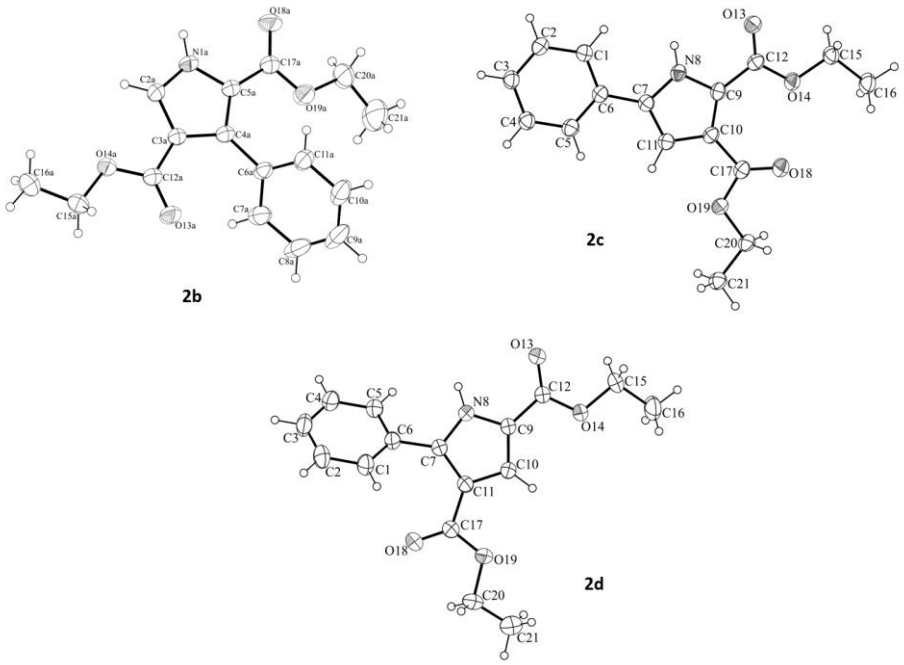


Figure 6. *The ORTEP drawing of phenyl-1H-pyrrole-diesters* ***2b-2d*** *with thermal ellipsoids at 30% level.*

In addition, the 3D spatial determinations of **1a**, **1d**, **1g** and **1j** were established by X-ray crystallography (***Figure 7***), and confirmed the structures in the solid state as anticipated on the basis of IR and in solution by ^{1}H and ^{13}C NMR data.

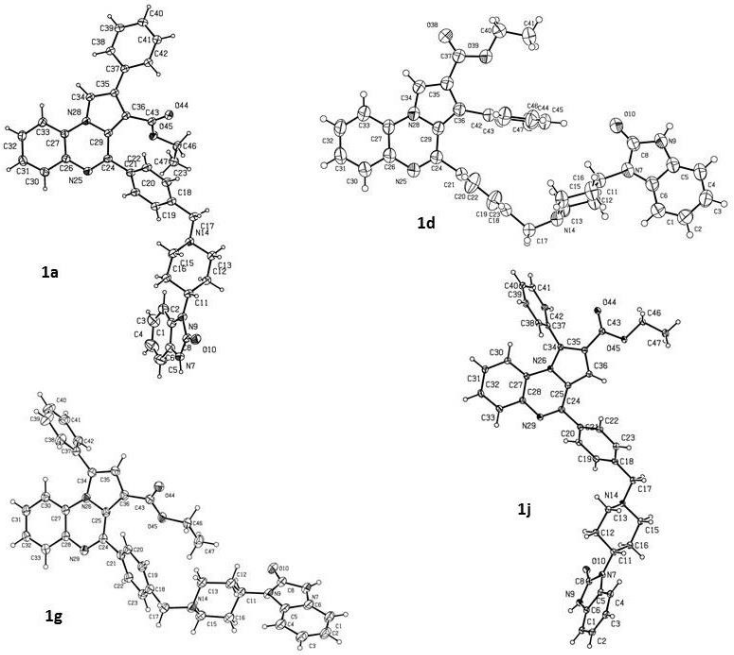


Figure7. *The ORTEP drawing of pyrrolo[1,2-a]quinoxalines* ***1a****,* ***1d****,* ***1g*** *and* ***1j*** *with thermal ellipsoids at 30% level.*

1.3.3. Biological activity

1.3.3.1. Cytotoxicity in different human cancer cell lines

The twelve new compounds **1a-l** were tested by MTS assay for their *in vitro* antiproliferative activity against five human leukemic cell lines (U937, K562, Jurkat, U266 and HL60). Compound A6730 (***Figure 4***) was used in these tests as the reference standard drug. The results are summarized in Table 2. The pyrrolo[1,2-*a*]quinoxalines **1c**, **1f**, **1g**, **1h, 1i, 1k** and **1l** were found the most antiproliferative compounds on the growth of human myeloid U937 cell line with IC_{50} of 3-11 μM. The two derivatives **1h** and **1k** with IC_{50} of 4 and 3 μM respectively, showed a better activity in comparison with the reference compound A6730 (IC_{50} = 8 μM). These two compounds **1h** and **1k** were substituted in position 4 by a benzylpiperidinyl fluorobenzimidazole group and in position 1 by a phenyl. In general, the substitution by a phenyl on position 1 of the pyrrole moiety led to active derivatives (**1g-l** excepted **1j**). Interestingly, **1c**, **1f, 1i** and **1l** were substituted by a benzylpiperidinyl triazolylpyridine moiety in position 4 of the pyrrolo[1,2-*a*]quinoxaline core. All other compounds derived from the incorporation of the benzylpiperidinyl benzimidazolone moiety (compounds **1a, 1d** and **1j**), which was present in the reference compound A6730, into the 4-position of the heterocyclic pyrroloquinoxaline ring were found inactive on the U937 cell line in comparison with their benzylpiperidinyl fluorobenzimidazole or benzylpiperidinyl triazolylpyridine analogues, excepted **1g** that was found active (IC_{50} = 9 μM). From a SAR point of view, these preliminary biological results on U937 cell line enlightened the importance of the substitution at C-4 position of the pyrroloquinoxaline scaffold by a benzylpiperidinyl fluorobenzimidazole group, and also the need of a phenyl functionalisation in position 1 of the pyrrole ring (compounds **1h** and **1k**). The antiproliferative potencies of these new derivatives **1a-l** were also examined towards the human myeloid leukaemia cell lines K562 and HL60. Among the twelve compounds tested for antiproliferative activities on K562 cell line, the five pyrrolo[1,2-*a*]quinoxalines **1a**, **1b**, **1e**, **1k** and **1l** were found the most active compounds with an IC_{50} of 3-4 μM. All the other pyrroloquinoxalines **1** also showed significant antiproliferative activity with IC_{50} ranging from 7 to 14 μM, better than the one found for the reference compound A6730 (IC_{50} = 17 μM). Moreover, in terms of structure-activity relationships discussion, it could be also noticed that the four quinoxalines **1b**, **1e**, **1h** and **1k**, bearing the benzylpiperidinyl fluorobenzimidazole moiety in their 4-position, were always found the most active compounds with IC_{50} of 3-7 μM in each subseries diversely substituted by a phenyl and an ester on the pyrrole ring. Against the HL60 human acute promyeloid leukemia cell line, most of the tested compounds showed antiproliferative activity with IC_{50} values from 3 to 24 μM, excepted **1a** and **1j** that were found inactive (IC_{50}> 50 μM). The two pyrroloquinoxalines **1d** and **1f** having an ester function and a phenyl respectively in position 2 and 3 exhibited better activities than their other homologues. Compounds **1d** and **1k** showed a better activity that the one noticed for the reference compound A6730: i.e. IC_{50} = 3.5 and 3 μM for **1d** and **1k**, respectively in comparison with 5.5 μM for A6730. The antiproliferative activities of compounds **1a-l** against the T-lymphocyte Jurkat cell line were also investigated, and the results exhibited potent cytotoxicity for pyrroloquinoxalines **1b-l** (IC_{50} from 2 to 6.5 μM) as potent as the one observed for A6730 (IC_{50} = 3.5 μM). Nevertheless, the pyrrolo[1,2-*a*]quinoxaline **1a** showed low antiproliferative activity (IC_{50} = 41 μM). Hence in the subseries functionalized by a phenyl and an ester at C2 and C3

respectively, the IC_{50} of **1b** and **1c** (5 μM) was 8.2 times lower than those of compound **1a** (IC_{50} = 41 μM). Against the human myeloma cell line U266, the same pyrrolo[1,2-*a*]quinoxalines **1b**, **1f**, **1h**, **1k** and **1l**, bearing a benzylpiperidinyl fluorobenzimidazole or a benzylpiperidinyl triazolylpyridine moiety in position 4 and substituted on the pyrrole ring, exhibited potent cytotoxicity (IC_{50} from 3 to 5 μM). All the other tested pyrroloquinoxalines **1** (compounds **1d-e**, **1g** and **1i**) were also found to be active on U266 cell line with IC_{50} ranging from 8 to 18 μM, with the exception of **1a**, **1c** and **1j** that presented an IC_{50} superior to 50 μM.

1.3.3.2. Cytotoxicity activity in activated normal PBMNC cells

The compounds **1a-l** were tested on activated (PBMNC + PHA) human peripheral blood mononuclear cells to evaluate their respective cytotoxicity on normal cells (***Table 2***). As expected, most of the pyrrolo[1,2-*a*]quinoxalines **1a-l** showed significant level of cytotoxicity against lymphocytes with IC_{50} ranging from 8 to over 50 μM. These preliminary results were used to determine their respective range of toxic concentration. Indexes of selectivity (IS) were defined as the ratio of the IC_{50} value on the human mononuclear cells to the IC_{50} value on the K562, U937, HL60, Jurkat and U266 lines. Compounds that demonstrated high selectivity (high index of selectivity) should offer a potential of safer therapy. This led to identify compounds with index of selectivity >16.7 and >12.5 for compounds **1e** and **1a**, respectively, on the human myeloid leukemic cell lines K562; and >12.5 for compound **1g** against the human leukemic cell lines Jurkat. We could notice that the more interesting new pyrroloquinoxaline structure (compound **1e**) could be considered as a direct combination of our previously described bioactive derivatives JG572 and JG576 (**Series B**, ***Figure 4***). Moreover, we could notice that the compound **1a** showed interesting selectivity towards K562 CML cell lines. The potential inhibitor **1h** also showed interesting IS on U937 and U266 leukemic cell lines with value of 12.5. These four compounds could now constitute suitable candidates for further pharmacological studies. The reference compound A6730 showed interesting selectivity with index of selectivity value noticed at 14.3 on the Jurkat cell line.

***Table 2**. In vitro activity of compounds **1a-l** (Series D, **Figure 3b**) on K562, U937, HL60, Jurkat and U266 cells, and cytotoxicity on human peripheral blood mononuclear cells PBMNC + PHA.*

IC_{50} values (μM)[a]						
Inhibitors	**K562**	**U937**	**HL60**	**Jurkat**	**U266**	**Cytotoxicity on activated human peripheral blood mononuclear cells (PBMNC) PBMNC + PHA**
A6730	17 ± 0.3	8 ± 0.2	5.5 ± 0.2	3.5 ± 0.2	n.d.[b]	>50
1a	4 ± 0.1	>50	>50	41 ± 1.2	>50	>50
1b	3 ± 0.1	21 ± 2.2	19 ± 2.8	5 ± 0.2	3.5 ± 0.1	14 ± 1
1c	14 ± 0.3	7 ± 0.3	12 ± 2.3	5 ± 0.1	>50	16 ± 2
1d	9 ± 0.3	>50	3.5 ± 0.1	4 ± 0.1	10 ± 0.8	41.3 ± 5
1e	3 ± 0.1	n.d.	10 ± 0.9	5 ± 0.2	18 ± 1.1	>50
1f	7 ± 0.2	7 ± 0.1	5 ± 0.1	2 ± 0.1	5 ± 0.1	8 ± 0.5
1g	8 ± 0.2	9 ± 0.4	6 ± 0.1	4 ± 0.1	9 ± 0.8	>50
1h	7 ± 0.3	4 ± 0.1	8 ± 0.2	6 ± 0.15	4 ± 0.1	50 ± 4
1i	12 ± 0.4	11 ± 0.9	24 ± 3	5 ± 0.1	8 ± 0.5	50 ± 6
1j	8.5 ± 0.3	>50	>50	6.5 ± 0.1	>50	>50
1k	3 ± 0.1	3 ± 0.1	3 ± 0.1	3.5 ± 0.1	3 ± 0.1	13 ± 1
1l	3.5 ± 0.1	8 ± 0.2	7 ± 0.3	3	3 ± 0.1	12 ± 1

[a]The IC_{50}(μM) values correspond to the mean ± standard deviation from three independent experiments.
[b]n.d. = not determined.

According to these results (***Table 2***), the molecules **1e**, **1g**, **1h** and **1i** exhibit the most promising activity on most of the cancer cell lines.

1.3.4. Prediction of toxicity and other drug relevant properties

To speculate about any toxicity risks and evaluate the drug-like characteristics of these novel synthesized compounds, we computed a set of drug relevant properties from their 2D chemical structures (***Table 3***). The calculated properties indicated that our molecules **1** present the same toxicological profile as the anticancer reference compound A6730. These parameters were calculated by the molinspiration web services. **[http://www.molinspiration.com/cgi-bin/properties. Accession date: July 20, 2015]** All compounds **1a-l** were found lipophilic with Clog *P* values between 6.44 and 7.70. The Clog *P* of the reference compound A6730 was noticed at 5.55 slightly lower than those of our compounds. To predict intestinal absorption, we estimated the molecular polar surface areas (PSA) of these new pyrroloquinoxaline compounds **1a-l** from the calculated TPSA. PSA has been extensively used in medicinal chemistry for modeling absorption phenomena and to optimize a drug's ability to permeate cells. The PSA of a molecule is defined as the surface sum over all polar atoms, which primarily consist of oxygen and nitrogen as well as their attached hydrogens. Molecules with a polar surface area greater than 140 $Å^2$ tend to be poor at permeating cell membranes, while a PSA of less than 60 $Å^2$ is usually needed for molecules to penetrate the blood-brain barrier and thus act on the brain and other central nervous system tissues [33,34]. The PSA of our compounds could suggest an intestinal absorption (PSA < 140 $Å^2$). On the other hand, these pyrroloquinoxalines seem to be unlikely to cross the blood-brain barrier (PSA > 60 $Å^2$). Thus, based on these predicted data, the calculated PSA values for these derivatives **1** could indicate a possible better oral bioavailability with a smaller chance of CNS toxicity.

Table 3*. Predicted drug-relevant properties of compounds* ***1a-l****.*

Compound	Clog *P*	TPSA	nON	nOH/NH	N violations[a]
A6730	5,55	95,50	8	2	2
1a	7.19	84.65	8	1	2
1b	7.70	75.53	7	1	2
1c	6.49	101.32	9	1	2
1d	7.19	84.65	8	1	2
1e	7.70	75.53	7	1	2
1f	6.49	101.32	9	1	2
1g	7.13	84.65	8	1	2
1h	7.65	75.53	7	1	2
1i	6.44	101.32	9	1	2
1j	7.13	84.65	8	1	2
1k	7.65	75.53	7	1	2
1l	6.44	101.32	9	1	2

[a]Number of violations to the Lipinski's "rule of five": log *P*≤5, molecular weight ≤500, number of hydrogen bond acceptors≤10, and number of hydrogen bond donors ≤5.

1.3.5. Docking studies

1.3.5.1. Validation of Swiss Dock server method

Firstly, docking study was carried out considering the reference molecule A6730 as ligand (***Figure 8a***). The best docked structure was compared to the X-ray structure of Akt1/A6730 (pdb code: 3O96 [5], **Figure 8b**) inhibitor complex. The superimposition of the two structures showed that A6730 molecule in the best docked pose has position and orientation almost identical to that assumed in the crystal structure (see ***Figure 8c***). Furthermore, structure-activity relationship (SAR) studies have underlined the importance of the 2,3-diphenyl substituent and the N1 heteroatom (blue in the ***Figure 8a***). All these results confirm what is reported in the reference work [5]. For this reason, they could be considered as a valuable evidence of the Swiss Dock server validity.

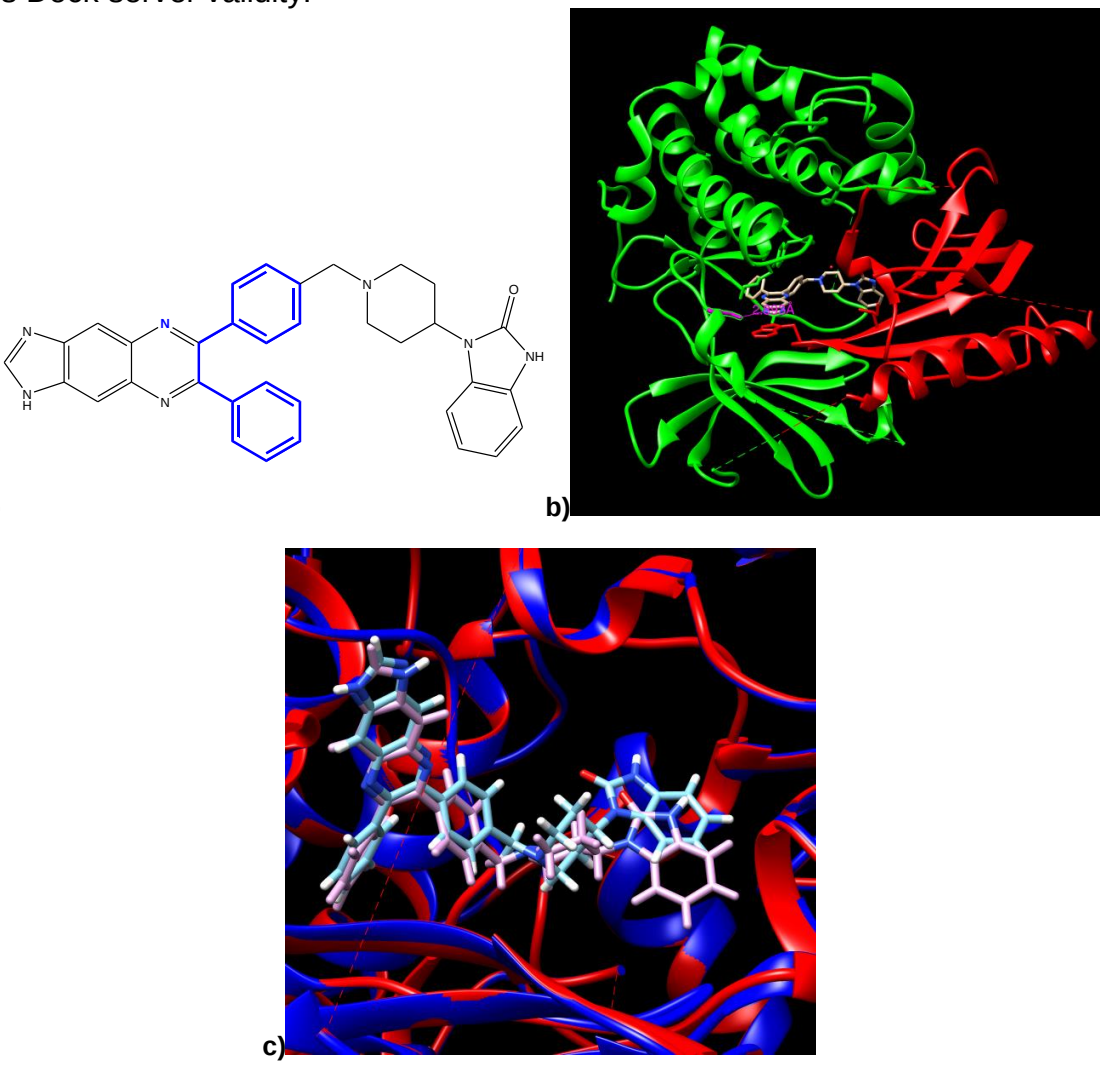


Figure 8. a) *Chemical structure of A6730 allosteric inhibitor; Blue the molecule portions involved in the Akt-ligand interaction.* ***b)*** *Reference complex: X-ray structure of Akt1/A6730 inhibitor complex (pdb code:3O96 [5]); Red for PH domain; Green for kinase domain.****c)*** *Superimposition of the reference complex (pdb: 3O96) and best docking pose of A6730; Red for reference Akt structure (pdb: 3O96); Blue for Akt structure used for docking analysis; Light blue for the ligand in the best docking pose of A6730; Pink for molecule A6730 in the reference complex (pdb code:3O96).*

1.3.5.2. Docking analysis of the compounds 1a-l with Akt entire protein

The results of docking studies conducted on our new synthesized molecules **1a-l** are presented by visualizing the two Akt domains (PH and Kinase domains) separately to give a better comprehension of the main protein-ligand interactions.

1.3.5.2.1. Compound 1a

The first molecule studied in this work is the pyrroloquinoxaline **1a** (***Figure 5***). The result of docking analysis consists of 256 models organized in 37 clusters. The first five clusters were considered to define the main interactions between ligand and protein. The ***Figure 9a*** shows the best docked structure (this represents the model with the lowest value of the FullFitness score) and in this figure only the Akt PH domain is visible. The protein residues making the highest number of contacts with the ligand are Gln79, Trp80 and Ile84. The ***Figure 9b*** contains the same docking pose but only the Kinase domain is represented. The residues with the highest number of contacts with the ligand are Tyr272 and Arg273. This model is characterized by the absence of intermolecular protein/ligand hydrogen bonds.

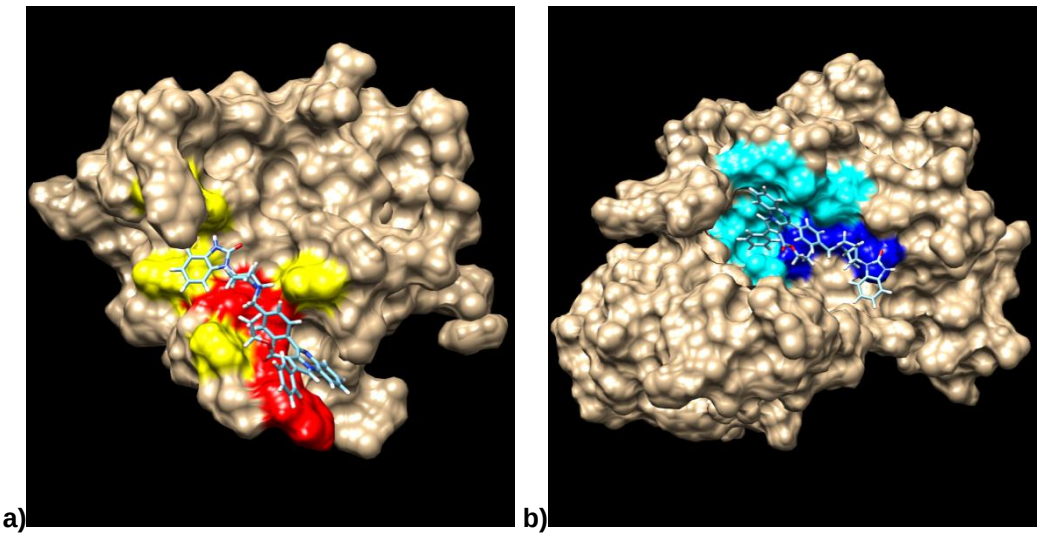


Figure 9. *Best docked structure for* ***1a****.* ***a)*** *only PH domain is shown. Red: residues providing 10 interactions or more; yellow for residues providing 9 interactions or less.* ***b)*** *only Kinase domain is reported. Blue for residues with 10 interactions or more. Cyan for residues with 9 interactions or less.*

1.3.5.2.2. Compound 1b

The second molecule studied is the derivative **1b** (***Figure 5***). The result of docking analysis consists of 256 models organized in 39 clusters. The first best scoring five clusters were analyzed to define the main interactions between ligand and protein. Most of the docking poses consist of the ligand fit in a cleft located at the interface between the Akt PH and Kinase domains.

The ***Figure 10a*** shows the best solution (this represents the model with the lowest value of the FullFitness score) and only the Akt PH domain is visible. The PH domain residues making the highest number of contacts with the ligand are Gln79 and Trp80. The ***Figure 10b*** contains the same docking pose but only the Kinase domain is represented. The residues belonging to the Kinase domain that are making the highest number of contacts with the ligand are Val270 and Tyr272. This docking model contains also a hydrogen bond between the backbone oxygen atom of Cys296 and the proton H31 of the benzimidazol portion of the ligand (***Figure 10c***).

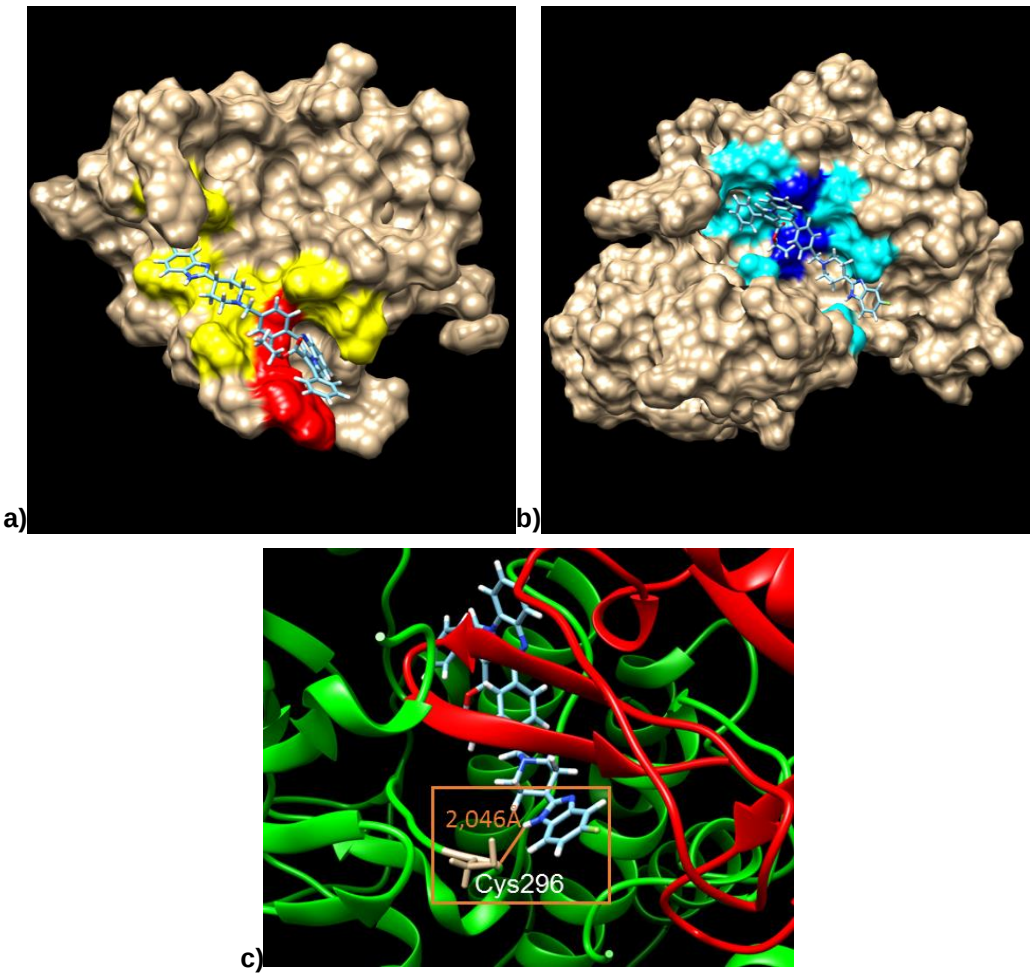


Figure 10. *Best docked structure for* ***1b****.* ***a)*** *only PH domain is shown. Red: residues providing 10 interactions or more; yellow for residues providing 9 interactionsor less.* ***b)*** *only Kinase domain. Blue for residues with 10 interactions or more. Cyan for residues with 9 interactions or less.* ***c)****Hydrogen bond analysis. Red for PH domain. Green for kinase domain. Gray for residue involved in the hydrogen bond. Orange for hydrogen bond length.*

1.3.5.2.3. Compound 1c

The third molecule studied is the compound **1c** (***Figure 5***). The result of docking analysis consists of 255 models organized in 36 clusters. The first best scoring five clusters were analyzed to define the main interactions between ligand and protein. Most of the docking poses consist of the ligand fit in a cleft located at the interface between the Akt PH and Kinase domains.

The ***Figure 11a*** shows the best solution (this represents the model with the lowest value of the FullFitness score) and only the Akt PH domain is visible. The PH domain residues making the highest number of contacts with the ligand are Lys14, Asn54, Gln79 and Trp80. The ***Figure 11b*** contains the same docking pose but only the Kinase domain is represented. The residues belonging to the Kinase domain that are making the highest number of contacts with the ligand are Tyr272, Asp292 and Tyr326. This docking model contains also a hydrogen bond between the HG1 proton of Thr82 and the O1 oxygen atom of the ethyl ester group of the ligand (***Figure 11c***).

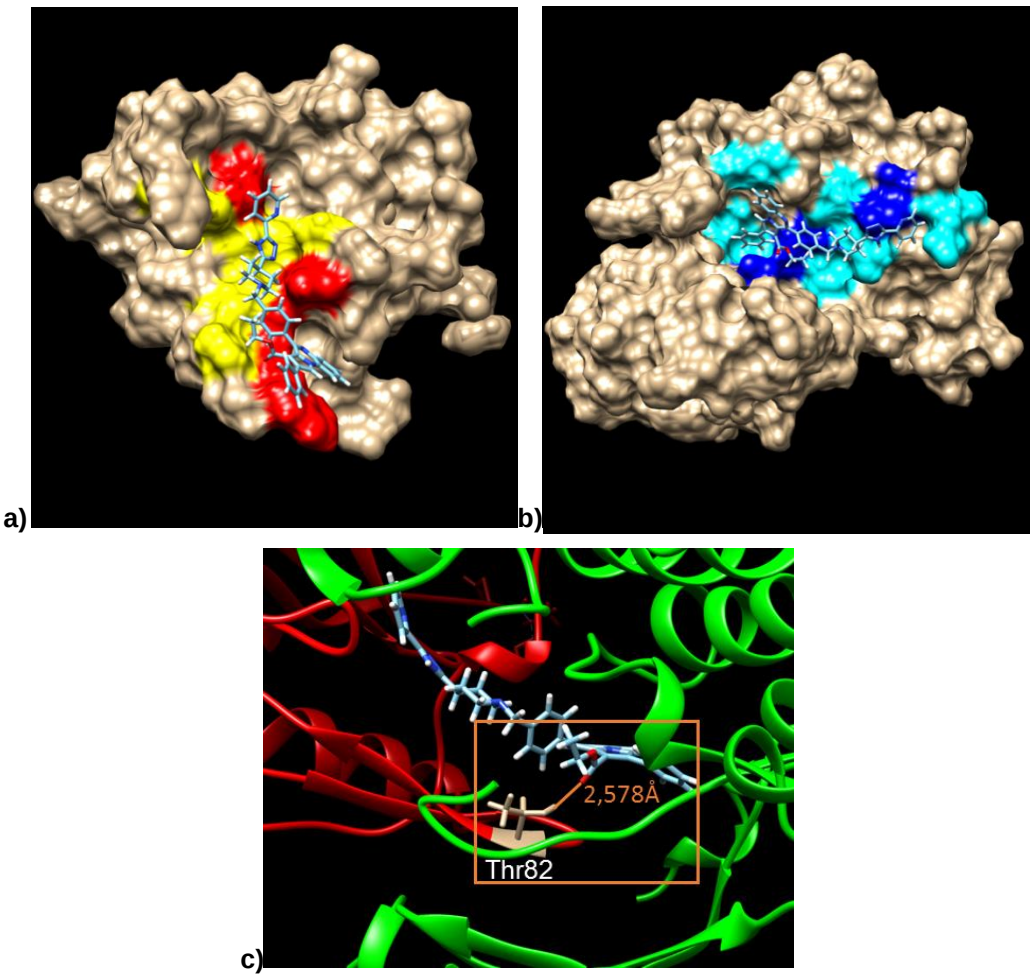


Figure 11. *Best docked structure for* ***1c****.* ***a)*** *only PH domain is shown. Red: residues providing 10 interactions or more; yellow for residues providing 9 interactionsor less.* ***b)*** *only Kinase domain. Blue for residues with 10 interactions or more. Cyan for residues with 9 interactions or less.* ***c)****Hydrogen bond analysis. Red for PH domain. Green for kinase domain. Gray for residue involved in the hydrogen bond. Orange for hydrogen bond length.*

1.3.5.2.4. Compound 1d

The fourth molecule studied is the pyrroloquinoxaline **1d** (***Figure 5***). The result of docking analysis consists of 254 models organized in 45 clusters. The first best scoring five clusters were analyzed to define the main interactions between ligand and protein. Most of the docking poses consist of the ligand fit in a cleft located at the interface between the Akt PH and Kinase domains.

The ***Figure 12a*** shows the best solution (this represents the model with the lowest value of the FullFitness score) and only the Akt PH domain is visible. The PH domain residues making the highest number of contacts with the ligand are Gln79 and Trp80.

The ***Figure 12b*** contains the same docking pose but only the Kinase domain is represented. The residues belonging to the Kinase domain that are making the highest number of contacts with the ligand are Val270 and Tyr272. This docking model contains also a hydrogen bond between the backbone H_N amide proton of Glu298 and the O atom of the benzimidazol-2-one portion of the ligand (***Figure 12c***).

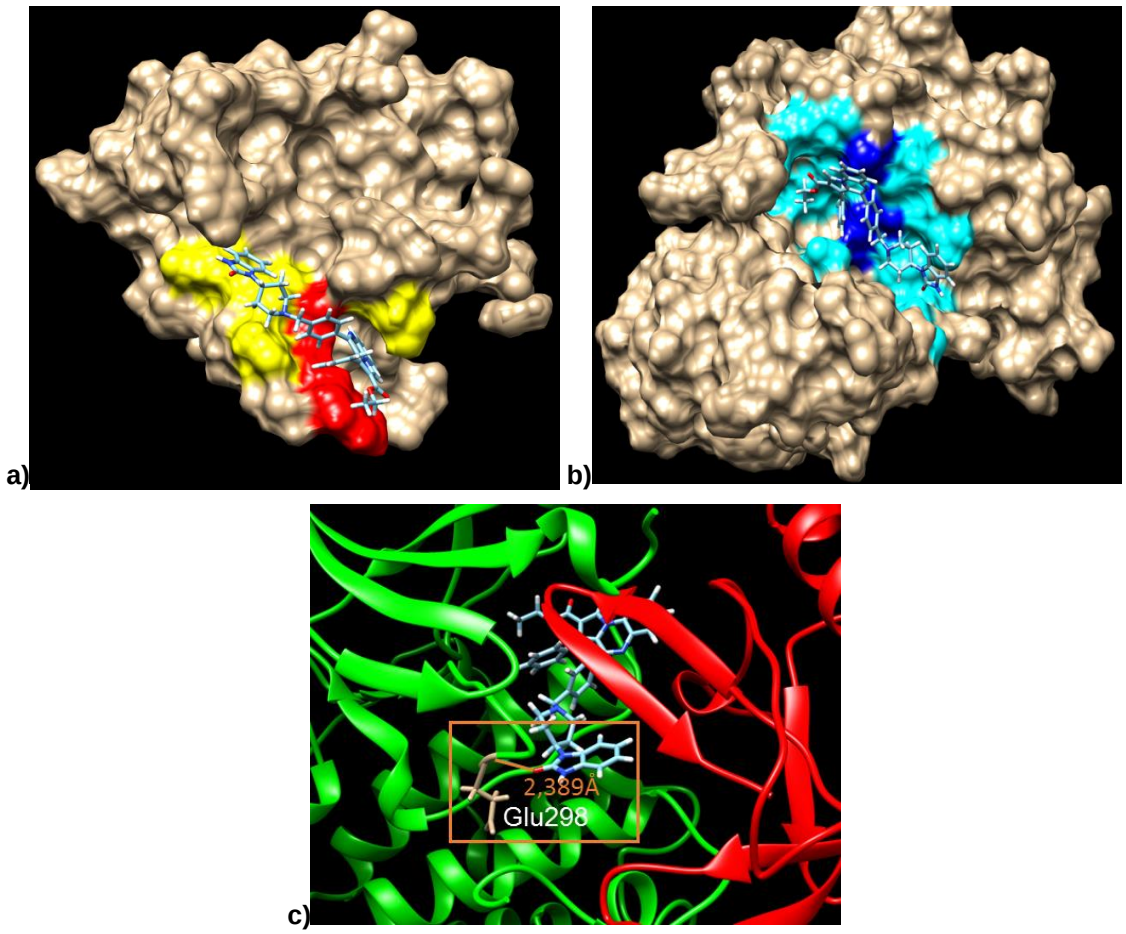


Figure 12. *Best docked structure for* ***1d****.* ***a)*** *only PH domain is shown. Red: residues providing 10 interactions or more; yellow for residues providing 9 interactionsor less.* ***b)*** *only Kinase domain. Blue for residues with 10 interactions or more. Cyan for residues with 9 interactions or less.* ***c)****Hydrogen bond analysis. Red for PH domain. Green for kinase domain. Gray for residue involved in the hydrogen bond. Orange for hydrogen bond length.*

1.3.5.2.5. Compound 1e

The fifth molecule studied is the molecule **1e** (***Figure 5***). The result of docking analysis consists of 256 models organized in 41 clusters. The first five clusters were investigated to identify the relevantinteractions between ligand and protein. In ***Figure 13a*** only the best model, provided with the lowest FullFitness score, is reported and the Akt PH domain is shown. The protein residues with the highest number of contacts with the ligand areGln79 and Trp80. The ***Figure 13b*** shows details of the corresponding docking pose in the Kinase domain cleft. Val270 and Tyr272 represent the residues of the Kinase domain making the highest number of contacts with the ligand. This model is characterized by the absence of intermolecular protein/ligand hydrogen bonds.

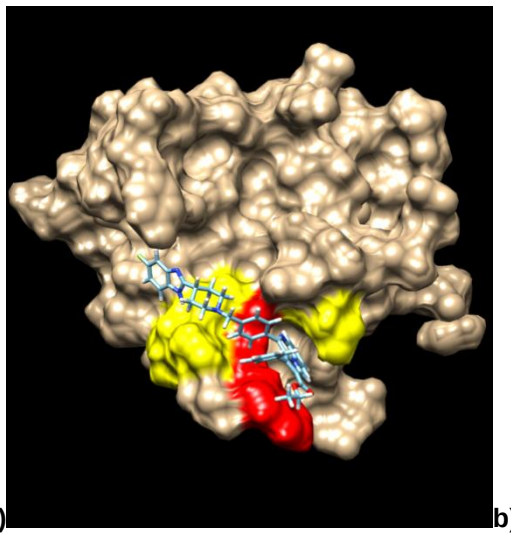
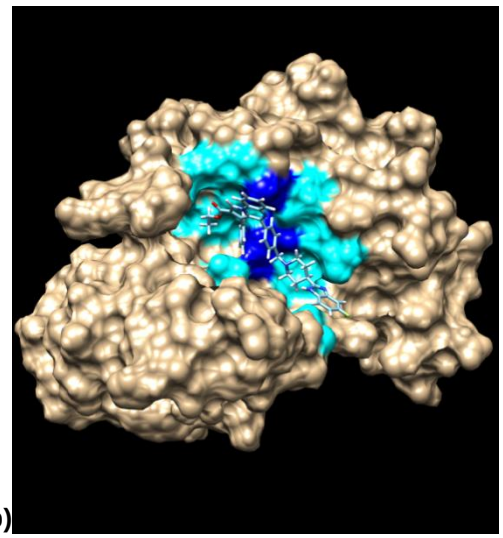
a)  b) 

Figure 13. *Best docked structure for* ***1e****.* ***a)*** *only PH domain is shown. Red: residues providing 10 interactions or more; yellow for residues providing 9 interactionsor less.* ***b)*** *only Kinase domain. Blue for residues with 10 interactions or more. Cyan for residues with 9 interactions or less.*

1.3.5.2.6. Compound 1f

The sixth molecule studied is the compound **1f** (***Figure 5***). The result of docking analysis consists of 256 models organized in 39 clusters. The first five clusters were considered to define the main interactions between ligand and protein. The ***Figure 14a*** shows the best docked structure (this represents the model with the lowest value of the FullFitness score) and in this figure only the Akt PH domain is visible. The protein residue making the highest number of contacts with the ligand is Trp80. The ***Figure 14b*** contains the same docking pose but only the Kinase domain is represented. The residues with the highest number of contacts with the ligand are Val270 and Tyr272. This model is characterized by the absence of hydrogen bonds with both domains.

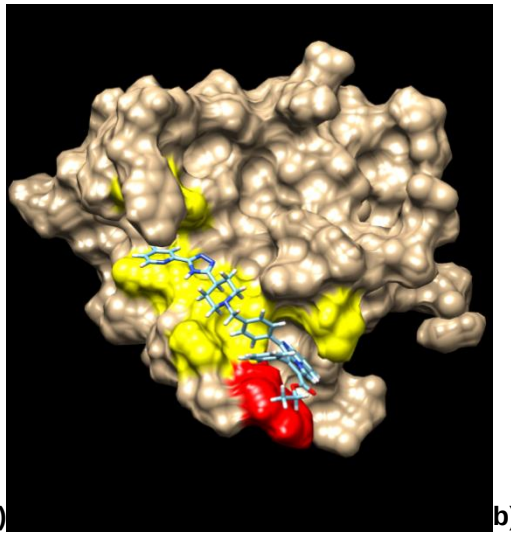
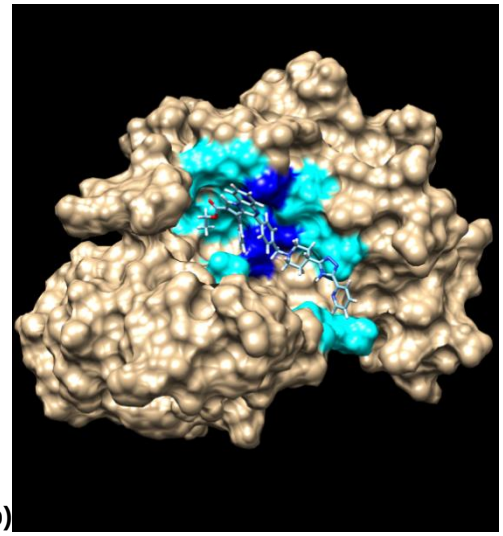
a)  b) 

Figure 14. *Best docked structure for* ***1f****.* ***a)*** *only PH domain is shown. Red: residues providing 10 interactions or more; yellow for residues providing 9 interactionsor less.* ***b)*** *only Kinase domain. Blue for residues with 10 interactions or more. Cyan for residues with 9 interactions or less.*

1.3.5.2.7. Compound 1g

The seventh molecule studied is pyrroloquinoxaline **1g** (***Figure 5***). The result of docking analysis consists of 256 models organized in 38 clusters. The first five clusters were considered to define the main interactions between ligand and protein. The ***Figure 15a*** shows the best docked structure (this represents the model with the lowest value of the FullFitness score) and in this figure only the Akt PH domain is visible. The protein residues making the highest number of contacts with the ligand are Gln79 and Trp80. The ***Figure 15b*** contains the same docking pose but only the Kinase domain is represented. The residues with the highest number of contacts with the ligand are Val270 and Arg273. This model is characterized by the absence of intermolecular protein/ligand hydrogen bonds.

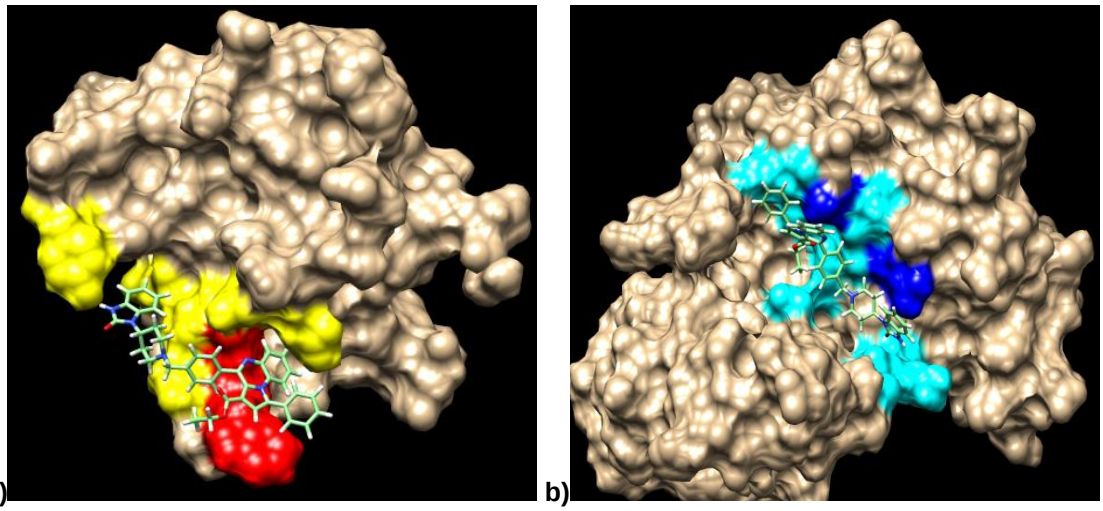


Figure 15. *Best docked structure for* ***1g****.* ***a)*** *only PH domain is shown. Red: residues providing 10 interactions or more; yellow for residues providing 9 interactionsor less.* ***b)*** *only Kinase domain. Blue for residues with 10 interactions or more. Cyan for residues with 9 interactions or less.*

1.3.5.2.8. Compound 1h

The eighth molecule studied is the derivative **1h** (***Figure 5***). The result of docking analysis consists of 256 models organized in 38 clusters. The first five clusters were considered to define the main interactions between ligand and protein. In the ***Figure 16a*** only the Akt PH domain is shown with the best model provided with the lowest FullFitness score. The protein residues with the highest number of contacts with the ligand are Gln79 and Trp80. The ***Figure 16b*** shows details of the corresponding model docked into the Kinase domain cleft. The residue making the highest number of contacts with the ligand is Val270. This model is characterized by the absence of hydrogen bonds with both domains.

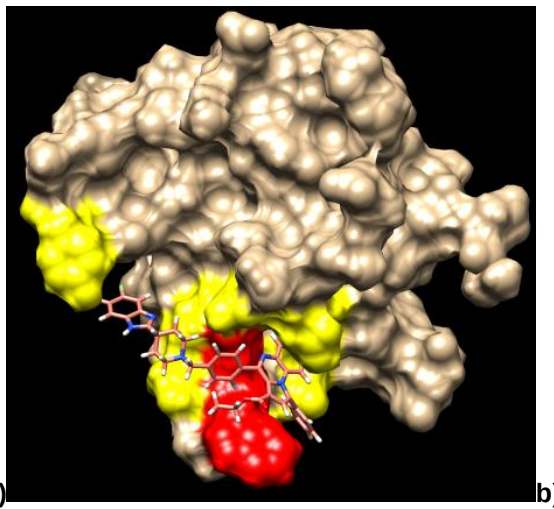
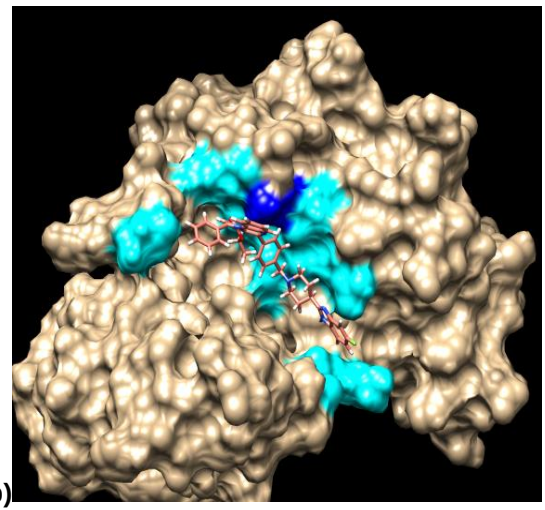
a)  b) 

Figure 16. *Best docked structure for* ***1h****.* ***a)*** *only PH domain is shown. Red: residues providing 10 interactions or more; yellow for residues providing 9 interactionsor less.* ***b)*** *only Kinase domain. Blue for residues with 10 interactions or more. Cyan for residues with 9 interactions or less.*

1.3.5.2.9. Compound 1i

The ninth molecule studied is the compound **1i** (***Figure 5***). The result of docking analysis consists of 252 models organized in 36 clusters. The first five clusters were considered to define the main interactions between ligand and protein. The ***Figure 17a*** shows the best docked structure (this represents the model with the lowest value of the FullFitness score) and in this figure only the Akt PH domain is visible. The protein residue making the highest number of contacts with the ligand is Trp80. The ***Figure 17b*** contains the same docking pose but only the Kinase domain is represented. The residues with the highest number of contacts with the ligand are Lys268 andVal270. This model is characterized by the absence of intermolecular protein/ligand hydrogen bonds.

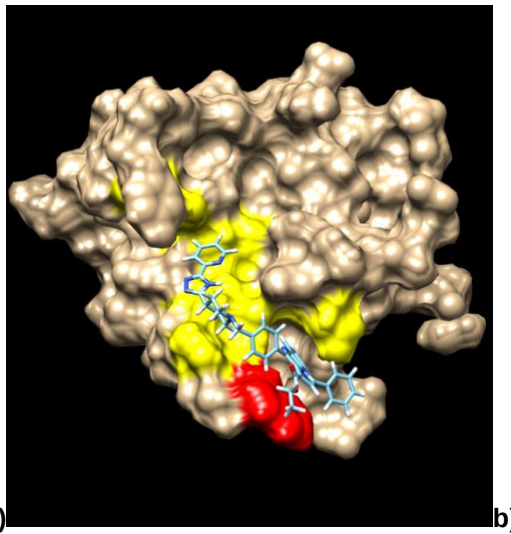
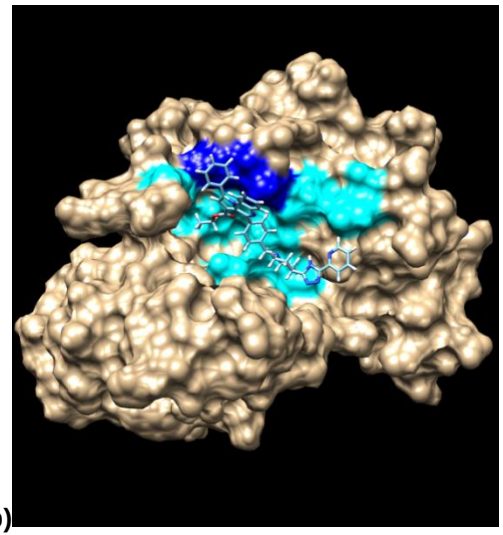
a)  b) 

Figure 17. *Best docked structure for* ***1i****.* ***a)*** *only PH domain is shown. Red: residues providing 10 interactions or more; yellow for residues providing 9 interactionsor less.* ***b)*** *only Kinase domain. Blue for residues with 10 interactions or more. Cyan for residues with 9 interactions or less.*

1.3.5.2.10. Compound 1j

The tenth molecule studied is pyrroloquinoxaline **1j** (***Figure 5***). The result of docking analysis consists of 256 models organized in 45 clusters. The first five clusters were considered to define the main interactions between ligand and protein. In the ***Figure 18a*** only the Akt PH domain is shown with the best model provided with the lowest FullFitness score. The protein residues with the highest number of contacts with the ligand are Asn53, Gln79 and Trp80. The ***Figure 18b*** shows details of the corresponding model docked into the Kinase domain cleft. The residue making the highest number of contacts with the ligand is Val270. This model is characterized by the absence of hydrogen bonds with both domains.

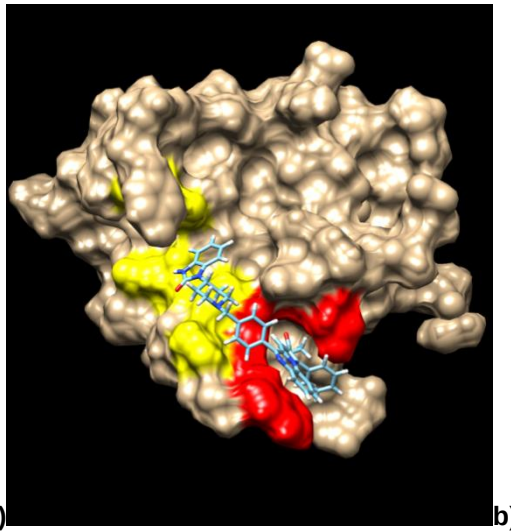
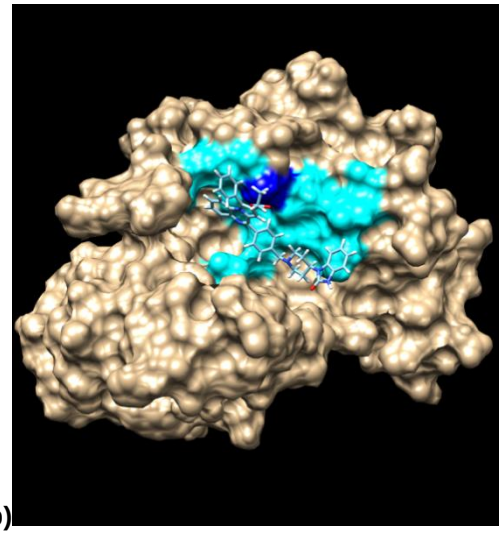
a)  **b)** 

Figure 18. *Best docked structure for **1j**. **a)** only PH domain is shown. Red: residues providing 10 interactions or more; yellow for residues providing 9 interactionsor less. **b)** only Kinase domain. Blue for residues with 10 interactions or more. Cyan for residues with 9 interactions or less.*

1.3.5.2.11. Compound 1k

The eleventh molecule studied is the molecule **1k** (***Figure 5***). The result of docking analysis consists of 256 models organized in 42 clusters. The first best scoring five clusters were analyzed to define the main interactions between ligand and protein. Most of the docking poses consist of the ligand fit in a cleft located at the interface between the Akt PH and Kinase domains.

The ***Figure 19a*** shows the best solution (this represents the model with the lowest value of the FullFitness score) and only the Akt PH domain is visible. The PH domain residue making the highest number of contacts with the ligand is Trp80. The ***Figure 19b*** contains the same docking pose but only the Kinase domain is represented. The residue belonging to the Kinase domain that is making the highest number of contacts with the ligand isLys268. This docking model contains also a hydrogen bond between the HD22 proton of Asn53 and the O1 oxygen atom of the ethyl ester group of the ligand (**Figure 19c**).

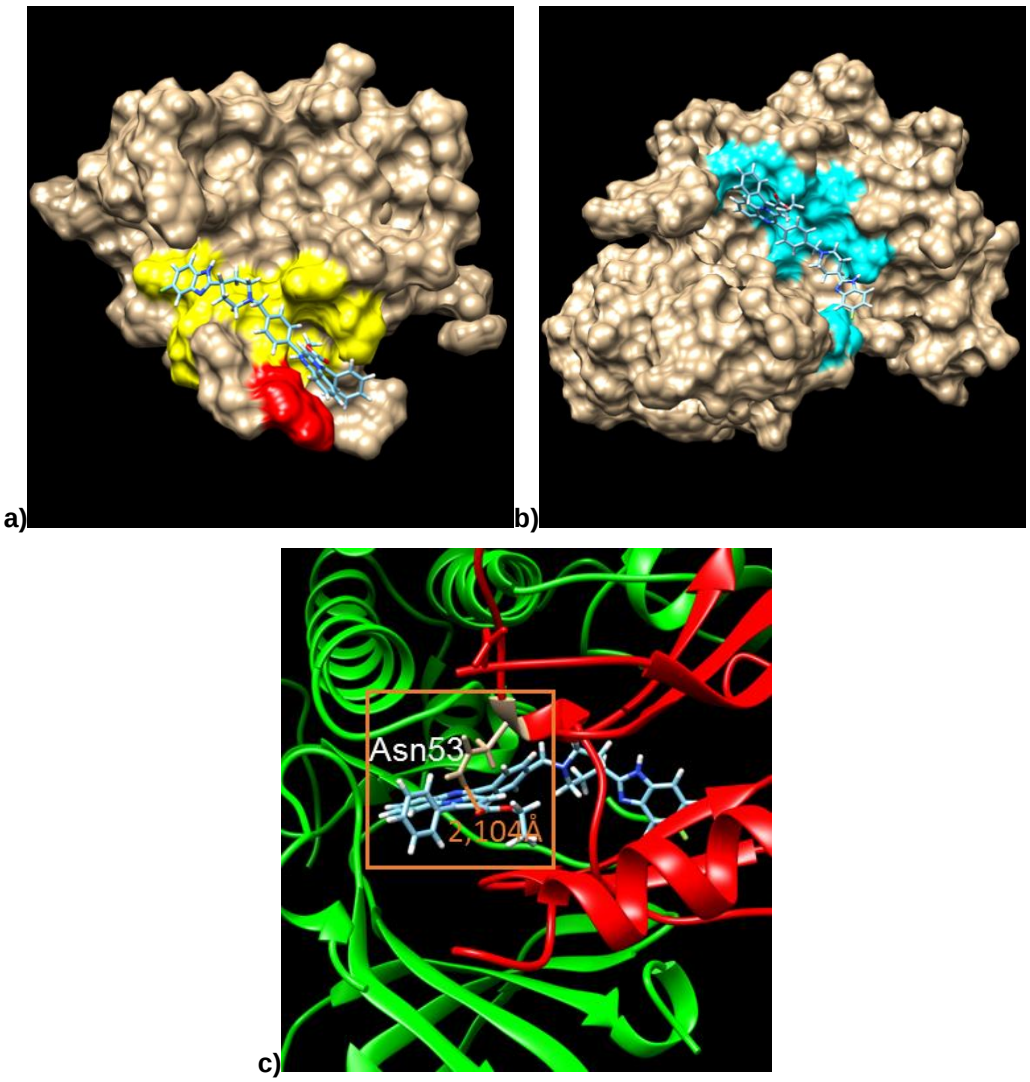


Figure 19. *Best docked structure for* ***1k****.* ***a)*** *only PH domain is shown. Red: residues providing 10 interactions or more; yellow for residues providing 9 interactionsor less.* ***b)*** *only Kinase domain. Blue for residues with 10 interactions or more. Cyan for residues with 9 interactions or less.* ***c)*** *Hydrogen bond analysis. Red for PH domain. Green for kinase domain. Gray for residue involved in the hydrogen bond. Orange for hydrogen bond length.*

1.3.5.2.12. Compound 1l

The last studied is the compound **1l** (***Figure 5***). The result of docking analysis consists of 255 models organized in 36 clusters. The first best scoring five clusters were analyzed to define the main interactions between ligand and protein. Most of the docking poses consist of the ligand fit in a cleft located at the interface between the Akt PH and Kinase domains. The ***Figure 20a*** shows the best solution (this represents the model with the lowest value of the FullFitness score) and only the Akt PH domain is visible. The PH domain residues making the highest number of contacts with the ligand are Asn53, Gln79 and Trp80. The ***Figure 20b*** contains the same docking pose but only the Kinase domain is represented. The residue belonging to the Kinase domain that is making the highest number of contacts with the ligand isVal270. This docking model contains also an hydrogen bond between the HH21 proton of Arg273 and the N4 nitrogen atom of the triazole portion of the ligand (***Figure 20c***).

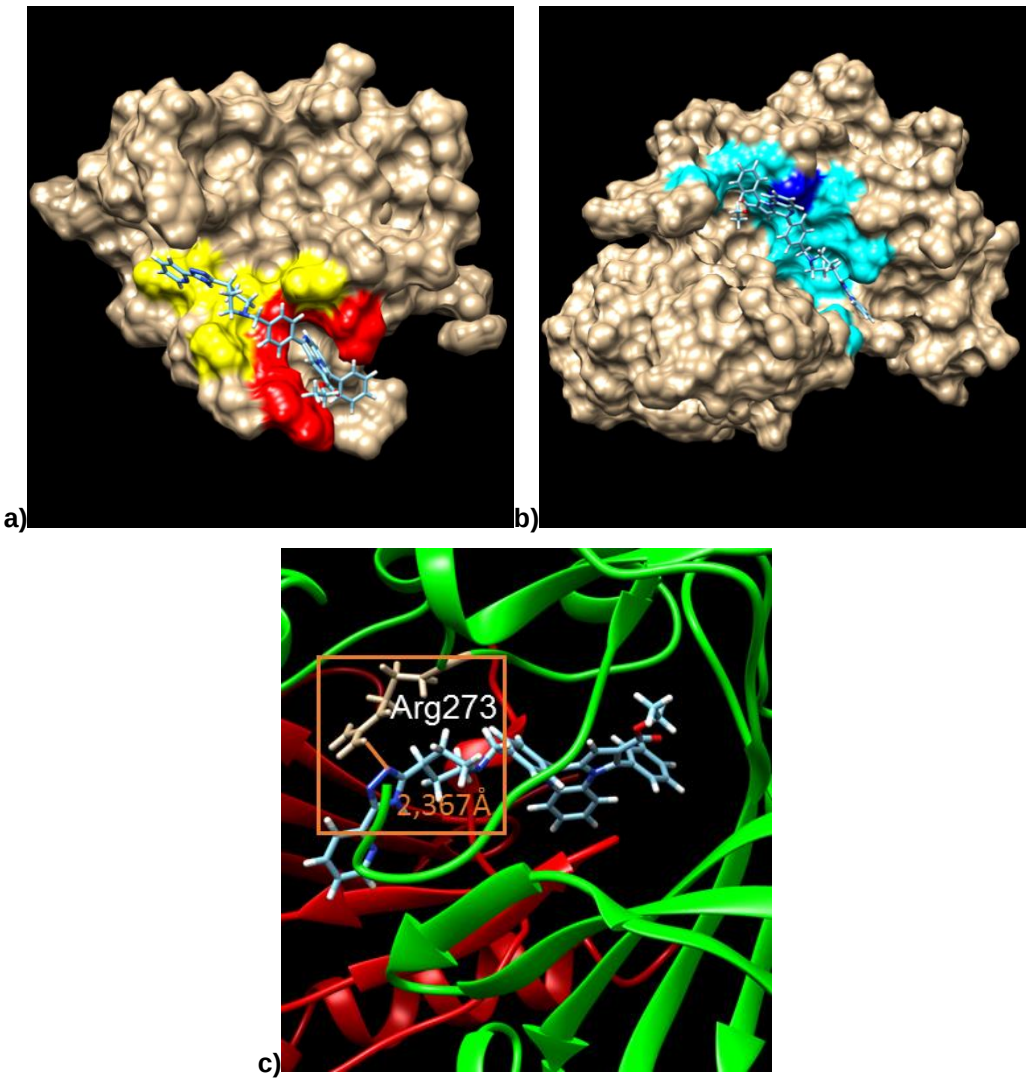


Figure 20. *Best docked structure for* ***1l****.* ***a)*** *only PH domain is shown. Red: residues providing 10 interactions or more; yellow for residues providing 9 interactionsor less.* ***b)*** *only Kinase domain. Blue for residues with 10 interactions or more. Cyan for residues with 9 interactions or less.* ***c)*** *Hydrogen bond analysis. Red for PH domain. Green for kinase domain. Gray for residue involved in the hydrogen bond. Orange for hydrogen bond length.*

1.3.5.2.13. Comparison between different ligands

All molecules fit pretty well in a cavity formed at the interface between the PH and Kinase domain; however, the molecule **1e** has the best (i.e. lowest) value of the FullFitness (-2350.714) (See ***Figures 13a-b***, ***Table 4***). This compound seems to fit slightly better than the other molecules in the cleft between the PH and Kinase domains. Thus, it is possible that the presence of a phenyl ring on the third position combined with the introduction of the small ethyl ester group on the second position of the pyrrolo[1,2-*a*]quinoxaline moiety could promote a better interaction between the ligand and the protein. The pyrrolo[1,2-*a*]quinoxaline **1h** appears as the worst ligand because has the highest value of the FullFitness score (-2328.7522) (See ***Figures 16a-b***, ***Table 4***). We could consider the hypothesis that the 5-fluorobenzimidazole ring, associated with a phenyl and an ester function substituted in positions 1 and 3 respectively, could interact with the residues of the binding pocket more difficultly.

***Table 4**. FullFitness values of the best docked structures for each molecule of Series D (**1a-l**).*

Compound	FullFitness
1a	-2343.2258
1b	-2346.2656
1c	-2332.023
1d	-2343.896
1e	-2350.714
1f	-2334.478
1g	-2341.46
1h	-2328.7522
1i	-2329.3647
1j	-2344.8442
1k	-2340.9988
1l	-2330.2043

1.3.5.3. Docking analysis of the compounds 1a-l with Akt PH domain

The results of docking studies with the PH domain reveal that **1e** is the best scoring compound. The best model obtained for this molecule is presented in ***Figure 21***. The protein residue with the highest number of contacts with the ligand isThr101.

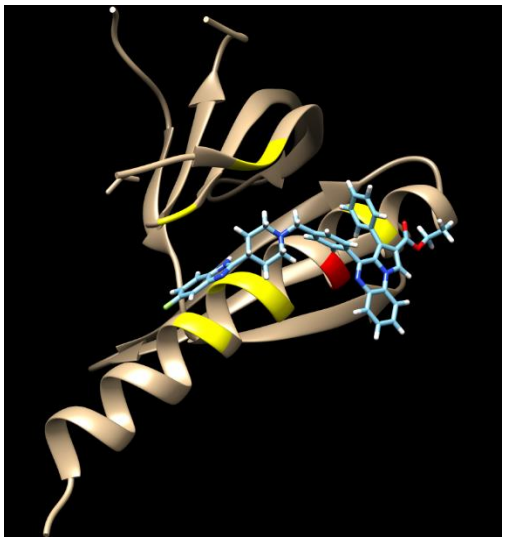


***Figure 21.** Best docked structure for **1e** considering only PH domain (pdb code: 1UNR [19]). Red: residues providing 10 interactions or more; yellow for residues providing 9 interactionsor less.*

The **1e** best docked structure obtained for Akt PH domain (***Figure 21***) was compared to a crystal structure of inositol-(1,3,4,5)-tetrakisphosphate (IP4) (***Figure 22a***) in complex with Akt1 PH domain (pdb code: 1UNQ [19]). The superimposition of the two complexes is reported in ***Figure 22b***.

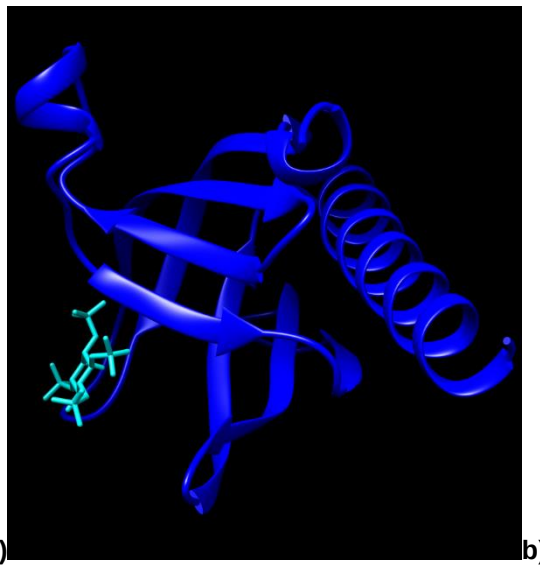
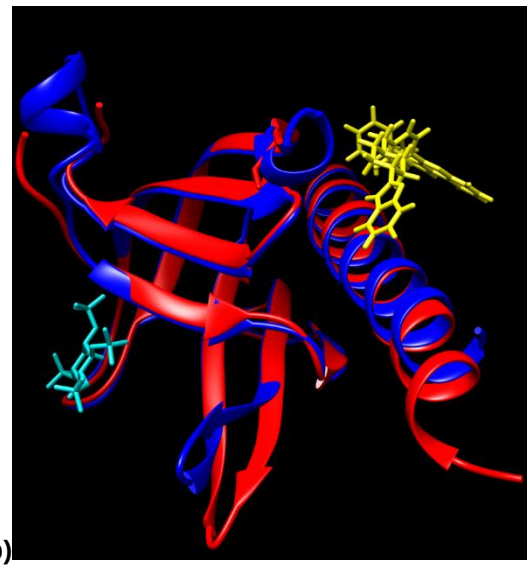
a)  b) 

Figure 22. a) *X-ray structure of Akt1 PH domain/IP4 complex (pdb code:1UNQ [19]) considered as reference. Cyan for the IP4 ligand. Blue for the Akt PH domain.* ***b)*** *Superimposition of best docking pose of* ***1e*** *molecule for PH domain (in red) and reference structure (in blue).*

Molecule **1e** fits on the PH domain surface in a pocket distinct from the inositol binding site (***Figures 22b***). In conclusion, docking results indicate that, more than single PH domain binders, the studied compound could behave as allosteric inhibitor able to interact with both the PH and kinase domains of Akt.

1.4. Conclusions

Attenuation of protein kinases by selective inhibitors is an extremely active field of activity in anticancer drug development. Therefore, Akt, a serine/threonine protein kinase, also known as protein kinase B (PKB), represents an attractive potential target for therapeutic intervention. AKT is a member of the serine/threonine AGC protein kinase family involved in different cellular processes such as metabolism, growth, proliferation and survival. Its three human isoenzymes (Akt1, Akt2 and Akt3) are highly homologous multi-domain proteins with both overlapping and distinct cellular functions (***Figure 1)***. Dysregulation of the AKT pathway has been identified in multiple human cancers. It is also well known that this protein is overexpressed in these cells. One of the possible consequences of this unusual protein expression is the uncontrolled proliferation of the cells. Therefore, the development of new potential antiproliferative molecules targeting this protein is required. Recent efforts in the design, synthesis and biological evaluation of small molecules, inhibitors of Akt, have led to the identification of novel inhibitors with various heterocycle scaffolds. Considering previous results obtained on the antiproliferative activities of new pyrrolo[1,2-*a*]quinoxalines (Series A-C, ***Figure 4***), a novel series (Series D, ***Figure 5***) was designed and synthesized from various substituted phenyl-1*H*-pyrrole-2-carboxylicacid alkyl esters *via* a multistep heterocyclization process. The antiproliferative activity of the twelve new synthesized ethyl 4-[4-(4-substitutedpiperidin-1-yl)]benzyl-phenylpyrrolo[1,2-a]quinoxaline-carboxylate derivatives **1a-l** was then investigated on the human leukemic cell lines U937, K562, Jurkat, U266 and HL60. Amond them, the four pyrrolo[1,2-a]quinoxalines **1a**, **1e**, **1g** and **1h** showed high cytotoxic activity against some cancer cell lines (IC_{50} ranging from 3 to 9 μM) and low toxicity against normal hematopoietic cells (estimated IC_{50} > 50 μM) (***Table 2***). These compounds, showing interesting index of selectivity, may constitute suitable candidates for further pharmacological studies. From a general SAR point of view, these preliminary biological results highlight the importance of substitution at the C-4 position of the pyrroloquinoxaline scaffold by a benzylpiperidinylfluorobenzimidazole group, and also the need for a functionalization on the pyrrole ring. Considering all biological results, the compound **1e** seem to be the most promising Akt ligand. Moreover, it would be also interesting to enlarge the biological evaluation of these new pyrrolo[1,2-*a*]quinoxaline derivatives in order to precise now their mechanism of action. Some drug relevant properties of the synthesized compounds **1a-l** were calculated by the molinspiration web services. All molecules exhibited the same toxicological profile of the anticancer reference compound A6730 (***Table 3***). One of the parameter considered was the hydrophilicity, represented by Clog *P* and correlated to the absorption or permeation of the molecules. Considering the range of Clog *P* values (6.44-7.70), the functioanalizations introduced in the pyrrolo[1,2-*a*]quinoxalines scaffold do not seem to improve the permeation of the molecules through the cell membrane. The molecular polar surface areas (PSA) has been extensively used in medicinal chemistry for modelling absorption phenomena and to optimize a drug's ability to permeate cells. The PSA values calculated for our compounds suggested an intestinal absorption (PSA < 140 $Å^2$). On the other hand, these molecules are unlikely to cross the blood-brain barrier and other central nervous system tissues (PSA > 60 $Å^2$). According these results, the synthesized molecules are characterized by a potential better oral bioavailability with a smaller chance of CNS toxicity. Docking studies confirmed biological results. In fact, the molecule **1e** exhibited the best (i.e. lowest) value of the FullFitness (-2350.714) (See ***Figures 13a-b***). Furthermore, it

seems to fit slightly better than the other molecules in the cleft between the PH and Kinase domains. This lead us to speculate that the presence of a phenyl ring on the third position combined with the introduction of the small ethyl ester group on the second position of the quinoxaline ring could promote a better interaction between the ligand and the protein. In fact, the molecule **1h**, characterized by the same functional groups but in different positions on the pyrrole ring, has the worst (i.e. highest) FullFitness score (-2328.7522) (See ***Figures 16a-b***). The docking results acquired during the study conducted on the isolated PH domain showed a similar trend. In particular, they reveal that **1e** is the best scoring compound. The superimposition of its best docked structure for Akt PH domain and the crystal structure of inositol-(1,3,4,5)-tetrakisphosphate (IP4) in complex with Akt1 PH domain (pdb code: 1UNQ [19]) (***Figure 22a***) highlights the presence of our ligand in a pocket distinct from the inositol binding site (**Figures 22b**). In conclusion docking results indicate that, more than single PH domain binders, the studied compound could be considered as an allosteric inhibitor for its ability to interact with both the PH and kinase domains of Akt. The possible toxicity of the molecules is a critical parameter to consider in the design of new potential drugs. In conclusion, a multidisciplinary and iterative protocol (synthesis, IR, crystallography, NMR, biological activities, docking and drug relevant properties predictions) was employed to obtain potential Akt ligands. This procedure could be repeated many times to improve the results as much as possible and to obtain new molecule suitable for *in vivo* studies. Very recently, in order to determine the possible mechanism of action of our compounds, we first evaluated their potency on isolated enzymes such as Akt 1, 2 and 3 as well as on mTOR. These tests have been performed by DiscoverX at 1 and 10 μM (KINOMEScan® / DiscoveRx, 11180 Roselle Street, SuiteD San Diego, CA 92121 United States). Nevertheless, only low effects have been detected. Akt2 activity was reduced of 38% and 35 % with derivatives **1e** and **1j** at 10 μM, respectively. Such of enzymatic interaction could not explain the cell effects. Next assessment of apoptosis in K562 cell line was examined. Apoptosis induces loss of membrane asymmetry resulting in phosphatidyl serine (PS) exposure and alterations in mitochondrial membrane potential. Cells were incubated with or without increasing doses **1e** for 3 days. Compound **1e**, which showed the most interesting index of selectivity on the human myeloid leukemic cell lines K562, induces a significantly increase in Annexin V positive cells from the second and third days (92% ± 2.8). These results could explain the inhibition of cell proliferation observed with this compound **1e**. Additional biological tests are now required to better define the mechanism of action of our compounds. These investigations will carry out at the University of Bordeaux by Dr. Vanessa Desplat.

1.5. Material and methods

1.5.1. Reactives

Commercially reagents were used as purchased without additional purification. Microwave experiments were carried out using a focused microwave reactor (CEM Discover).

1.5.2. Procedures

1.5.2.1. *Ethyl 2-azidoacrylate (4)*

To a solution of sodium azide (3.75 g, 57.7 mmol) in DMF (120 mL) at 65 °C was added ethyl 2,3-dibromopropionate (5.00 g, 19.3 mmol). After 10 min, the reaction mixture was cooled and poured into water (300 mL) and extracted with ether (3 x 120 mL). The combined organic extracts were washed with water (3 x 120 mL), dried over $MgSO_4$, filtered and evaporated in vacuo to afford ethyl 2-azidoacrylate as a yellow oil (79%). R_f=0.38 (cyclohexane/Et_2O-98/2) [37].

1.5.2.2. *Diethyl 4-phenyl-1H-pyrrole-2,3-dicarboxylate (2a)*

Conventional heating: To a 1,4-dioxane solution (10.0 mL) of 1,3-bis(diphenylphosphino)propane (dppp) (0.320 g, 0.780 mmol) were added ethyl isocyanoacetate (0.700 g, 6.19 mmol) and ethyl phenylpropiolate (0.900 g, 5.17 mmol). The solution was stirred at 100 °C for 2 h. After the consumption of ethyl isocyanoacetate, the reaction mixture was cooled to room temperature, filtered and evaporated *in vacuo*. The crude was purified on silica gel (eluent: cyclohexane/AcOEt-90/10 then cyclohexane/AcOEt-70/30, R_f=0.33) to afford **2a** as a colorless oil (87%). *Microwave heating:* To a 1,4-dioxane solution (18.0 mL) of 1,3-bis(diphenylphosphino)propane (dppp) (0.640 g, 1.56 mmol) were added ethyl isocyanoacetate (1.40 g, 13.4 mmol) and ethyl phenylpropiolate (1.80 g, 10.3 mmol). After a pre-stirring of 30 s, the solution was irradiated during 2 h. The irradiation was programmed to maintain a constant temperature (100 °C) with a power of 200 W. The reaction mixture was cooled to room temperature, filtered and concentrated *in vacuo*. The crude was purified on silica gel as above to give **2a** (88%) [20; 21].

1.5.2.3. *Diethyl 3-phenyl-1H-pyrrole-2,4-dicarboxylate (2b)*

Method A: To a mixture of ethyl phenylpropiolate (2.50 g, 14.3 mmol) and Ag_2CO_3 (0.264 g, 0.960 mmol) in 1,4-dioxane (40.0 mL) at 80 °C, ethyl isocyanoacetate (1.08 g, 9.55 mmol) was slowly added. The solution was stirred at 100 °C for 30 min. After the consumption of ethyl isocyanoacetate, the reaction mixture was cooled to room temperature, filtered and evaporated *in vacuo*. The crude was purified on silica gel (eluent: cyclohexane/AcOEt-90/10 then cyclohexane/AcOEt-70/30, R_f=0.33) to afford **2b** as white crystals (9%). Method B: *Conventional heating:* A mixture of ethyl phenylpropiolate (2.50 g, 14.3 mmol) and Ag_2CO_3 (0.264 g, 0.960 mmol) in 1,4-dioxane (40.0 mL) was heated at 80 °C for 5 min, then mixture was cooled with an ice bath. To this resulting solution cooled at room temperature, ethyl isocyanoacetate (1.08 g, 9.55 mmol) was added dropwise. The reaction mixture was stirred for 5 min at 25 °C, then heated for 30 min at 80 °C. The resulting slurry was concentrated under reduced pressure and taken up with dichloromethane. The organic layer was washed with brine, dried with Na_2SO_4, filtered and evaporated under *vacuum*. The crude was purified on silica gel as above leading to **2b** (10%). *Microwave heating:* To a 1,4-dioxane suspension (40.0 mL) of Ag_2CO_3 (0.264 g, 0.960 mmol) was added

ethyl phenylpropiolate (2.50 g, 14.3 mmol). After a pre-stirring of 30 s, the solution was irradiated during 1 min. The irradiation was programmed to maintain a constant temperature (80 °C) with a power of 200 W. To the resulting solution cooled at room temperature, ethyl isocyanoacetate (1.08 g, 9.55 mmol) was added dropwise. After a pre-stirring of 30 s, the solution was irradiated during 6 min with a first step of irradiation of 1 min at 25 °C followed by a second one of 5 min at 80 °C. The power was set at 200 W in both steps. The resulting slurry was concentrated under reduced pressure and taken up with dichloromethane. The organic layer was washed with brine, dried with Na_2SO_4, filtered and evaporated under *vacuum*. The crude was purified on silica gel as above to give **2b** (7%).Method C: To a *N*-methyl-2-pyrrolidone (NMP) solution (40.0 mL) of Ag_2CO_3 (0.264 g, 0.960 mmol) was added ethyl phenylpropiolate (2.50 g, 14.3 mmol). After a prestirring of 30 s, the solution was irradiated during 1 min. The irradiation was programmed to maintain a constant temperature (80 °C) with a power of 200 W. To the resulting solution cooled at room temperature, ethyl isocyanoacetate (1.08 g, 9.55 mmol) was added dropwise. After a pre-stirring of 30 s, the solution was irradiated during 6 min with a first step of irradiation of 1 min at 25 °C followed by a second one of 5 min at 80 °C. The power was set at 200 W in both steps. The resulting slurry was concentrated under reduced pressure and taken up with dichloromethane. The organic layer was washed with brine, dried with Na_2SO_4, filtered and evaporated under *vacuum*. The crude was purified on silica gel as above to afford **2b** (12%). Method D: *Conventional heating:* To a 1,4-dioxane solution (25.0 mL) of Cu_2O (0.0850 g, 0.590 mmol) and 1,10-phenanthroline (0.215 g, 1.19 mmol) were added ethyl isocyanoacetate (1.35 g, 11.9 mmol) and ethyl phenylpropiolate (2.50 g, 14.3 mmol). The solution was stirred at 100 °C for 2 h. After the consumption of ethyl isocyanoacetate, the reaction mixture was cooled to room temperature, filtered and evaporated *in vacuo*. The crude was purified on silica gel as above to give **2b** (58%). *Microwave heating:* To a 1,4-dioxane solution (25.0 mL) of Cu_2O (0.0850 g, 0.590 mmol) and 1,10-phenanthroline (0.215 g, 1.19 mmol) were added ethyl isocyanoacetate (1.35 g, 11.9 mmol) and ethyl phenylpropiolate (2.50 g, 14.3 mmol). After a prestirring of 30 s, the solution was irradiated during 40 min. The irradiation was programmed to maintain a constant temperature (100 °C) with a power of 200 W. The reaction mixture was cooled to room temperature, filtered and concentrated *in vacuo*. The crude was purified on silica gel as above to give **2b** (62%). Method E: *Conventional heating:* To a mixture of ethyl isocyanoacetate (1.98 g, 17.5 mmol) and 1,8-diazabicyclo[5.4.0]undec-7-ene (DBU) (3.00 g, 19.7 mmol) dissolved in THF (30.0 mL) was added dropwise benzaldehyde (1.06 g, 9.99 mmol) in THF (10.0 mL) at 50 °C for a period of 15 min with stirring. After stirring for 5 h at same temperature, the reaction mixture was neutralized with acetic acid and then the solvent was removed under reduced pressure. The resulting residue was extracted with ethyl acetate and the extract was washed with hydrochloric acid and water, dried with Na_2SO_4, and then evaporated *in vacuo*. The crude was purified on silica gel as above to give **2b** (21%). *Microwave heating:* To a THF solution (30.0 mL) of 1,8-diazabicyclo[5.4.0]undec-7-ene (DBU) (3.00 g, 19.7 mmol) was added ethyl isocyanoacetate (1.98 g, 17.5 mmol). After a prestirring of 30 s, the solution was irradiated during 2 min. The irradiation was programmed to maintain a constant temperature (50 °C) with a power of 200 W. To the resulting solution at the same temperature, ethyl isocyanoacetate (1.08 g, 9.55 mmol) was added dropwise. After a prestirring of 30 s, the solution was irradiated during 1 h. The irradiation was programmed to maintain a constant temperature (50 °C) with a power of 200 W. The reaction mixture was neutralized with acetic acid and

then the solvent was removed under reduced pressure. The resulting residue was extracted with ethyl acetate and the extract was washed with hydrochloric acid and water, dried with Na_2SO_4, and then evaporated *in vacuo*. The crude was purified on silica gel as above to give **2b** (35%) [22, 23, 26].

1.5.2.4. *Diethyl 5-phenyl-1H-pyrrole-2,3-dicarboxylate (2c)*

Conventional heating: To 1,4-diazabicyclo[2.2.2]octane (DABCO) (0.0830 g, 0.740 mmol) and (*E*)-acetophenone oxime **3** (1.00 g, 7.40 mmol) in dry toluene (20.0 mL) was added diethyl acetylenedicarboxylate (1.26 g, 7.40 mmol). The reaction mixture was heated for 6 min at 80 °C, then for 2 h at 170 °C under pressure. The resulting residue was cooled to room temperature, filtered and concentrated *in vacuo*. The crude was purified on silica gel (eluent: EP/AcOEt-80/20, R_f=0.27) to give **2c** as pink crystals (28%).*Microwave heating:* To 1,4-diazabicyclo[2.2.2]octane (0.0830 g, 0.740 mmol) and oxime **3** (1.00 g, 7.40 mmol) in dry toluene (20.0 mL) was added diethyl acetylenedicarboxylate (1.26 g, 7.40 mmol), and the resultant mixture was subjected to the two-stage microwave irradiation sequence (stage 1: 80 °C, 6 min; stage 2: 170 °C, 45 min). The power was set at 200 W in both steps. The reaction mixture was cooled to room temperature, filtered and concentrated *in vacuo*. The crude was purified on silica gel as above to give **2c** (28%) [29, 30].

1.5.2.5. *Diethyl 5-phenyl-1H-pyrrole-2,4-dicarboxylate (2d)*

To a solution of ethyl 2-azidoacrylate **4** (1.70 g, 12.1 mmol) and ethyl phenylacetoacetate (3.47 g, 18.0 mmol) in MeOH (30.0 mL) was added AcOH (1.45 g, 24.1 mmol) and manganese(III) acetate dihydrate (1.29 g, 4.81 mmol), and the solution was stirred at 40 °C for 2 h. The reaction mixture was quenched with pH 9 ammonium buffer ($AcONH_4$ and NH_4OH), and then extracted twice with AcOEt. The combined organic extracts were washed with brine, dried with Na_2SO_4, filtered and concentrated *in vacuo*. The resulting residue was cooled, triturated in Et_2O and filtered. The crystals formed were filtered, and dried under reduced pressure to give **2d** as white crystals (42%) [31].

1.5.2.6. *General procedure for diethyl 1-(2-nitrophényl)-phenyl-pyrrole-dicarboxylate (5a-d)*

Conventional heating: To a solution of diethyl phenyl-1*H*-pyrrole-dicarboxylate **2a-d** (2.75 mmol) in 11 mL of DMF was added cesium carbonate (3.30 mmol). The mixture was stirred at room temperature for 10 min, then 2-fluoro-nitrobenzene (4.12 mmol) was added. The reaction mixture was refluxed for 1 h 30 (15 h for **5d**), then was diluted in AcOEt (35.0 mL). The organic layer was washed with water (2 x 30.0 mL), then brine (30.0 mL) and dried over sodium sulfate. The organic layer was concentrated under vacuo to give a brown oil. After triturating in Et_2O a solid was obtained and filtered off, washed with Et_2O and dried to give the desired product **5**. *Microwave heating:* A suspension of diethyl phenyl-1*H*-pyrrole-dicarboxylate **2a-d** (3.40 mmol), 1-fluoro-2-nitrobenzene (5.10 mmol) and cesium carbonate (4.06 mmol) in 12.0 mL of DMF was irradiated during 10 min. The irradiation was programmed to maintain a constant temperature (150 °C) with a maximal power output of 200 W. The reaction mixture was then diluted in AcOEt (60.0 mL), washed with water (2 x 50.0 mL), then brine (50.0 mL) and dried over sodium sulphate. The organic layer was concentrated under vacuo to give products **5a-d** as an oil.

1.5.2.6.1. *Ethyl 1-(2-nitrophenyl)-4-phenyl-pyrrole-2,3-dicarboxylate (**5a**).* Orange oil (85% / 89%). ^{1}H NMR ($CDCl_3$) δ: 8.13 (dd, 1H, *J*=7.90 and 1.35 Hz, H-3'), 7.74 (ddd, 1H, *J*=7.90, 7.90 and 1.35 Hz, H-4'), 7.63 (ddd, 1H, *J*=7.90, 7.90 and 1.35 Hz, H-5'), 7.55-7.43 (m, 3H, H-6', H-2 phenyl and H-6 phenyl), 7.40-7.28 (m, 3H, H-3 phenyl, H-4 phenyl and H-5 phenyl), 7.02 (s, 1H, H-5), 4.39-4.30 (m, 2H, OCH_2), 4.12 (q, 2H, *J*=7.20 Hz, OCH_2), 1.31 (t, 3H, *J*=7.20 Hz, CH_3), 1.15 (t, 3H, *J*=7.20 Hz, CH_3). HRMS-ESI m/z $[M+Na]^+$ Calcd for $C_{22}H_{20}N_2O_6Na$: 431.1219, Found: 431.1224.

1.5.2.6.2. *Ethyl 1-(2-nitrophenyl)-3-phenyl-pyrrole-2,4-dicarboxylate (**5b**).* Yellow oil (89% / 95%). ^{1}H NMR ($CDCl_3$) δ: 8.19 (dd, 1H, *J*=7.80 and 1.50 Hz, H-3'), 7.77 (ddd, 1H, *J*=7.80, 7.80 and 1.50 Hz, H-4'), 7.67 (ddd, 1H, *J*=7.80, 7.80 and 1.50 Hz, H-5'), 7.59 (s, 1H, H-5), 7.51 (dd, 1H, *J*=7.80 and 1.50 Hz, H-6'), 7.39-7.35 (m, 5H, 5H phenyl), 4.14 (q, 2H, *J*=7.20 Hz, OCH_2), 3.89-3.80 (m, 2H, OCH_2), 1.13 (t, 3H, *J*=7.20 Hz, CH_3), 0.77 (t, 3H, *J*=7.20 Hz, CH_3). HRMS-ESI m/z $[M+Na]^+$ Calcd for $C_{22}H_{20}N_2O_6Na$: 431.1219, Found: 431.1236.

1.5.2.6.3. *Ethyl 1-(2-nitrophenyl)-5-phenyl-pyrrole-2,3-dicarboxylate (**5c**).* Yellow oil (65% / 76%). ^{1}H NMR ($CDCl_3$) δ: 8.05-8.01 (m, 2H, H-3' and H-6'), 7.62-7.56 (m, 2H, H-2 phenyl and H-6 phenyl), 7.31 (t, 1H, *J*=7.50 Hz, H-4'), 7.25-7.04 (m, 4H, H-5', H-3 phenyl, H-4 phenyl and H-5 phenyl), 6.77 (s, 1H, H-4), 4.34 (q, 2H, *J*=6.90 Hz, OCH_2), 4.12 (q, 2H, *J*=6.90 Hz, OCH_2), 1.38-1.14 (m, 6H, $2CH_3$). HRMS-ESI m/z $[M+Na]^+$ Calcd for $C_{22}H_{20}N_2O_6Na$: 431.1219, Found: 431.1248.

1.5.2.6.4. *Ethyl 1-(2-nitrophenyl)-5-phenyl-pyrrole-2,4-dicarboxylate(**5d**).* Yellow oil (28%). ^{1}H NMR ($CDCl_3$) δ: 8.04 (dd, 1H, *J*=7.50 and 1.80 Hz, H-3'), 7.66 (s, 1H, H-3), 7.52 (ddd, 1H, *J*=7.50, 7.50 and 1.80 Hz, H-4'), 7.46 (ddd, 1H, *J*=7.50, 7.50 and 1.80 Hz, H-5'), 7.26-7.16 (m, 6H, H-6' and 5H phenyl), 4.17 (q, 2H, *J*=6.90 Hz, OCH_2), 4.15 (q, 2H, *J*=6.90 Hz, OCH_2), 1.27 (t, 3H, *J*=6.90 Hz, CH_3), 1.18 (t, 3H, *J*=6.90 Hz, CH_3). HRMS-ESI m/z $[M+Na]^+$ Calcd for $C_{22}H_{20}N_2O_6Na$: 431.1219, Found: 431.1240.

1.5.2.7. *General procedure for ethyl 4,5-dihydro-4-oxo-5H-phénylpyrrolo[1,2-a]quinoxaline-carboxylate (6a-d)*

A suspension of **5** (2.60 mmol) and iron powder (9.55 mmol) in 11.0 mL of acetic acid was heated under reflux for 2 h. The reaction mixture was cooled, suspended in 20.0 mL of a 1 M aqueous solution of HCl, agitated, then filtered off, washed with HCl 1 M (8.00 mL), water, Et_2O and dried to give **6** as a fluffy white solid.

1.5.2.7.1. *Ethyl 4,5-dihydro-4-oxo-5H-2-phenylpyrrolo[1,2-a]quinoxaline-3-carboxylate (**6a**).* Beige crystals (69%), mp 240°C. IR (KBr) 3200, 2750 (NH), 1720 (COO), 1660 (CON). ^{1}H NMR (DMSO-d_6) δ: 11.49 (s, 1H, NH), 8.58 (s, 1H, H-1), 8.17 (d, 1H, *J*=8.10 Hz, H-9), 7.55 (d, 2H, *J*=7.50 Hz, H-2 phenyl and H-6 phenyl), 7.45 (t, 2H, *J*=7.50 Hz, H-3 phenyl and H-5 phenyl), 7.34-7.27 (m, 4H, H-4 phenyl, H-6, H-7 and H-8), 4.30 (q, 2H, *J*=7.20 Hz, OCH_2), 1.25 (t, 3H, *J*=7.20 Hz, CH_3). ^{13}C-NMR (DMSO-d_6) δ: 157.0 (C-4), 155.2 (C-7), 129.7 (C-9a), 122.6 (C-5a), 117.7 (C-1), 116.8 (C-3a), 116.0 (C-6), 112.3 (C-3), 111.0 (C-2), 108.9 (C-9), 100.8 (C-8), 55.4 (CH_3O). Anal. Calcd. for $C_{12}H_{10}N_2O_2$: C, 67.28; H, 4.71; N, 13.08. Found: C, 67.40; H, 4.81; N, 12.94.

1.5.2.7.2. *Ethyl 4,5-dihydro-4-oxo-5H-3-phenylpyrrolo[1,2-a]quinoxaline-2-carboxylate (**6b**)*.Beige crystals (88%), mp>310°C. IR (KBr) 3200, 2800 (NH), 1720 (COO), 1665 (CON). ^{1}H NMR (DMSO-d_6) δ: 11.25 (s, 1H, NH), 8.80 (s, 1H, H-1), 8.28 (d, 1H, *J*=7.80 Hz, H-9), 7.34-7.27 (m, 7H, H-6, H-8 and 5H phenyl,), 7.21 (t, 1H, *J*=7.80 Hz, H-7), 4.09 (q, 2H, *J*=6.90 Hz, OCH_2), 1.09 (t, 3H, *J*=6.90 Hz, CH_3). ^{13}C-NMR (DMSO-d_6) δ: 157.0 (C-4), 155.2 (C-7), 129.7 (C-9a), 122.6 (C-5a), 117.7 (C-1), 116.8 (C-3a), 116.0 (C-6), 112.3 (C-3), 111.0 (C-2), 108.9 (C-9), 100.8 (C-8), 55.4 (CH_3O). Anal. Calcd. for $C_{12}H_{10}N_2O_2$: C, 67.35; H, 4.78; N, 13.15. Found: C, 67.47; H, 4.92; N, 13.30.

1.5.2.7.3. *Ethyl 4,5-dihydro-4-oxo-5H-1-phenylpyrrolo[1,2-a]quinoxaline-3-carboxylate (**6c**)*. White crystals (58%), mp=219°C. IR (KBr) 3200, 2700 (NH), 1700 (COO), 1670 (CON). ^{1}H NMR (DMSO-d_6) δ: 11.50 (s, 1H, NH), 7.56-7.53 (m, 5H, 5H phenyl), 7.32 (d, 1H, *J*=7.70 Hz, H-9), 7.24 (t, 1H, *J*=7.70 Hz, H-8), 6.97 (d, 1H, *J*=7.70 Hz, H-6), 6.84 (t, 1H, *J*=7.70 Hz, H-7), 6.81 (s, 1H, H-2), 4.27 (q, 2H, *J*=6.90 Hz, OCH_2), 1.30 (t, 3H, *J*=6.90 Hz, CH_3). Anal. Calcd. for $C_{12}H_{10}N_2O_2$: C, 67.21; H, 4.64; N, 13.01. Found: C, 67.33; H, 4.74; N, 12.76.

1.5.2.7.4. *Ethyl 4,5-dihydro-4-oxo-5H-1-phenylpyrrolo[1,2-a]quinoxaline-2-carboxylate (**6d**)*. White crystals (84%), mp=230°C. IR (KBr) 3200, 2700 (NH), 1695 (COO), 1665 (CON). ^{1}H NMR (DMSO-d_6) δ: 11.53 (s, 1H, NH), 7.61-7.53 (m, 3H, 3H phenyl), 7.50-7.47 (m, 2H, 2H phenyl), 7.44 (s, 1H, H-3), 7.29 (d, 1H, *J*=7.90 Hz, H-6), 7.20 (t, 1H, *J*=7.90 Hz, H-8), 6.75 (t, 1H, *J*=7.90 Hz, H-7), 6.64 (d, 1H, *J*=7.90 Hz, H-9), 4.02 (q, 2H, *J*=6.90 Hz, OCH_2), 1.30 (t, 3H, *J*=6.90 Hz, CH_3). Anal. Calcd. for $C_{12}H_{10}N_2O_2$: C, 67.31; H, 4.74; N, 13.11. Found: C, 67.43; H, 4.95; N, 12.86.

1.5.2.8. *General procedure for ethyl 4-chloro-phénylpyrrolo[1,2-a]quinoxaline-carboxylate (8a-d)*

A solution of 5*H*-pyrrolo[1,2-a]quinoxalin-4-one **6** (4.00 mmol) in $POCl_3$ (8.00 mL) was refluxed for 2 h. After removing excess of reactive under vacuum, the residue was carefully dissolved in water at 0°C and the resulting solution was made basic with sodium carbonate. The precipitate was filtered, dried and recrystallized from ethyl acetate to give **7**.

1.5.2.8.1. *Ethyl 4-chloro-2-phenylpyrrolo[1,2-a]quinoxaline-3-carboxylate (**8a**)*. White crystals (77%), mp=110°C. IR (KBr) 1715 (COO). ^{1}H NMR ($CDCl_3$) δ: 7.97 (s, 1H, H-1), 7.87 (dd, 1H, *J*=8.10 and 1.20 Hz, H-9), 7.80 (dd, 1H, *J*=8.10 and 1.20 Hz, H-6), 7.58-7.32 (m, 7H, H-7, H-8 and 5H phenyl), 4.41 (q, 2H, *J*=7.20 Hz, OCH_2), 1.34 (t, 3H, *J*=7.20 Hz, CH_3). HRMS-ESI m/z $[M+H]^+$ Calcd for $C_{20}H_{15}N_2O_2ClNa$: 351.0900, Found: 351.0918.

1.5.2.8.2. *Ethyl 4-chloro-3-phenylpyrrolo[1,2-a]quinoxaline-2-carboxylate (**8b**)*. Beige crystals (96%), mp=145°C. IR (KBr) 1715 (COO). ^{1}H NMR ($CDCl_3$) δ: 8.56 (s, 1H, H-1), 7.93 (dd, 1H, *J*=8.10 and 1.50 Hz, H-9), 7.88 (dd, 1H, *J*=8.10 and 1.50 Hz, H-6), 7.59 (ddd, 1H, *J*=8.10, 7.90 and 1.50 Hz, H-8), 7.52 (ddd, 1H, *J*=8.10, 7.90 and 1.50 Hz, H-7), 7.44-7.37 (m, 5H, 5H phenyl), 4.19 (q, 2H, *J*=7.20 Hz, OCH_2), 1.14 (t, 3H, *J*=7.20 Hz, CH_3). HRMS-ESI m/z $[M+H]^+$ Calcd for $C_{20}H_{15}N_2O_2ClNa$: 351.0900, Found: 351.0933.

1.5.2.8.3. *Ethyl 4-chloro-1-phenylpyrrolo[1,2-a]quinoxaline-3-carboxylate (**8c**)*. Beige crystals (94%), mp=109°C. IR (KBr) 1720 (COO). ^{1}H NMR ($CDCl_3$) δ: 7.91 (d, 1H, *J*=8.10 Hz, H-9), 7.55-7.48 (m, 5H, 5H phenyl), 7.41 (t, 1H, *J*=8.10 Hz, H-8), 7.33 (d, 1H, *J*=8.10 Hz, H-6), 7.17 (t, 1H, *J*=8.10 Hz, H-7), 7.09 (s, 1H, H-2), 4.44 (q, 2H, *J*=7.20 Hz, OCH_2), 1.44 (t, 3H, *J*=7.20 Hz, CH_3). HRMS-ESI m/z $[M+H]^+$ Calcd for $C_{20}H_{15}N_2O_2ClNa$: 351.0900, Found: 351.0903.

1.5.2.8.4. *Ethyl 4-chloro-1-phenylpyrrolo[1,2-a]quinoxaline-2-carboxylate (**8d**)*. Orange crystals (90%), mp=178°C. IR (KBr) 1710 (COO). ^{1}H NMR ($CDCl_3$) δ: 7.86 (dd, 1H, *J*=8.10 and 1.30 Hz, H-6), 7.63-7.55 (m, 4H, 3H phenyl and H-3), 7.50-7.46 (m, 2H, 2H phenyl), 7.37 (ddd, 1H, *J*=8.10, 7.85 and 1.30 Hz, H-8), 7.10 (ddd, 1H, *J*=8.10, 7.85 and 1.30 Hz, H-7), 7.01 (dd, 1H, *J*=8.10 and 1.30 Hz, H-9), 4.19 (q, 2H, *J*=7.20 Hz, OCH_2), 1.18 (t, 3H, *J*=7.20 Hz, CH_3). HRMS-ESI m/z $[M+H]^+$ Calcd for $C_{20}H_{15}N_2O_2ClNa$: 351.0900, Found: 351.0925.

1.5.2.9. *General procedure for ethyl 4-(4-formylphenyl)-phenylpyrrolo[1,2-a]quinoxaline-carboxylate (9a-d)*

To suspension of compound **8a-d** (4.64 mmol), and $Pd(PPh_3)_4$ (0.232 mmol) in a mixture of toluene/EtOH (75.0/4.10 mL) under nitrogen were added K_2CO_3 (5.10 mmol) and 4-formylphenylboronic acid (5.10 mmol). The reaction mixture was refluxed for 24 h, and the cooled suspension was extracted with CH_2Cl_2 (3 x 80.0 mL). The organic layer was washed with a saturated solution of NaCl (95.0 mL), and the combined organic extracts were dried over sodium sulfate, filtered, and evaporated under reduced pressure. The crude residue was triturated in ethanol. The resulting precipitate was filtered, washed with ethanol, and purified by column chromatography on silica gel using dichloromethane as eluent gave the pure product **9a-d**.

1.5.2.9.1. *Ethyl 4-(4-formylphenyl)-2-phenylpyrrolo[1,2-a]quinoxaline-3-carboxylate (**9a**)*. Yellow crystals (83%), mp=77°C. IR (KBr) 1720 (COO), 1700 (CHO). ^{1}H NMR ($CDCl_3$) δ: 10.13 (s, 1H, CHO), 8.12 (dd, 1H, *J*=8.10 and 1.30 Hz, H-9), 8.10 (s, 1H, H-1), 8.04 (d, 2H, *J*=7.80 Hz, H-3' and H-5'), 8.00 (dd, 1H, *J*=8.10 and 1.30 Hz, H-6), 7.96 (d, 2H, *J*=7.80 Hz, H-2' and H-6'), 7.67 (ddd, 1H, *J*=8.10, 7.85 and 1.30 Hz, H-8), 7.64-7.52 (m, 3H, H-7 and 2H phenyl), 7.48-7.39 (m, 3H, 3H phenyl), 3.63 (q, 2H, *J*=6.90 Hz, OCH_2), 0.89 (t, 3H, *J*=6.90 Hz, CH_3). HRMS-ESI m/z $[M+H]^+$ Calcd for $C_{27}H_{20}N_2O_3Na$: 421.1552, Found: 421.1552.

1.5.2.9.2. *Ethyl 4-(4-formylphenyl)-3-phenylpyrrolo[1,2-a]quinoxaline-2-carboxylate (**9b**)*. Yellow crystals (60%), mp=161°C. IR (KBr) 1715 (COO), 1700 (CHO). ^{1}H NMR ($CDCl_3$) δ: 9.93 (s, 1H, CHO), 8.69 (s, 1H, H-1), 8.09 (d, 1H, *J*=7.80 Hz, H-9), 8.05 (d, 1H, *J*=7.80 Hz, H-6), 7.66 (t, 1H, *J*=7.80 Hz, H-8), 7.58 (t, 1H, *J*=7.80 Hz, H-7), 7.53 (d, 2H, *J*=7.20 Hz, H-3' and H-5'), 7.38 (d, 2H, *J*=7.20 Hz, H-2' and H-6'), 7.03-6.95 (m, 5H, 5H phenyl), 4.23 (q, 2H, *J*=6.90 Hz, OCH_2), 1.20 (t, 3H, *J*=6.90 Hz, CH_3). HRMS-ESI m/z $[M+H]^+$ Calcd for $C_{27}H_{20}N_2O_3Na$: 421.1552, Found: 421.1532.

1.5.2.9.3. *Ethyl 4-(4-formylphenyl)-1-phenylpyrrolo[1,2-a]quinoxaline-3-carboxylate (**9c**)*. Pale yellow crystals (79%), mp=179°C. IR (KBr) 1720 (COO), 1695 (CHO). ^{1}H NMR ($CDCl_3$) δ: 10.14 (s, 1H, CHO), 8.09-8.06 (m, 1H, H-9), 8.04 (d, 2H, *J*=8.10 Hz, H-3' and H-5'), 7.97 (d, 2H, *J*=8.10 Hz, H-2' and H-6'), 7.58-7.56 (m, 5H,

5H phenyl), 7.48-7.44 (m, 2H, H-6 and H-8), 7.25-7.20 (m, 1H,H-7), 7.21 (s, 1H, H-2), 3.80 (q, 2H, *J*=7.20 Hz, OCH_2), 0.99 (t, 3H, *J*=7.20 Hz, CH_3). HRMS-ESI m/z $[M+H]^+$ Calcd for $C_{27}H_{20}N_2O_3Na$: 421.1552, Found: 421.1573.

1.5.2.9.4. *Ethyl 4-(4-formylphenyl)-1-phenylpyrrolo[1,2-a]quinoxaline-2-carboxylate (**9d**)*. Yellow crystals (89%), mp=183°C. IR (KBr) 1700 (COO and CHO). ^1H NMR ($CDCl_3$) δ: 10.18 (s, 1H, CHO), 8.20 (d, 2H, *J*=7.95 Hz, H-3' and H-5'), 8.12 (d, 2H, *J*=7.95 Hz, H-2' and H-6'), 8.02 (d, 1H, *J*=7.80 Hz, H-6), 7.62-7.48 (m, 6H, 5H phenyl and H-3), 7.42 (t, 1H, *J*=7.80 Hz, H-8), 7.13-7.09 (m, 2H, H-7 and H-9), 4.16 (q, 2H, *J*=7.20 Hz, OCH_2), 1.13 (t, 3H, *J*=7.20 Hz, CH_3). HRMS-ESI m/z $[M+H]^+$ Calcd for $C_{27}H_{20}N_2O_3Na$: 421.1552, Found: 421.1566.

1.5.2.10. *General procedure for ethyl 4-{4-[(4-(2-oxo-2,3-dihydro-1H-benzimidazol-1-yl)piperidin-1-yl)benzyl]}-phenylpyrrolo[1,2-a]quinoxaline-carboxylate, ethyl 4-{4-[(4-(5-fluoro-1H-benzimidazol-2-yl)piperidin-1-yl)benzyl]}-phenylpyrrolo[1,2-a]quinoxaline-carboxylate and ethyl 4-{4-[4-(3-(pyridin-2-yl)-1,2,4-triazol-5-yl)piperidin-1-yl)benzyl]}-phenylpyrrolo[1,2-a]quinoxaline-carboxylate (1a-l)*

The pH of a solution of the aldehyde **9a-d** (0.784 mmol) and 4-(2-ketobenzimidazolin-1-yl)piperidine or 4-(5-chloro-2-ketobenzimidazolin-1-yl)piperidine or 2-(3-piperidin-4-yl-1*H*-1,2,4-triazol-5-yl)pyridine (0.941 mmol) in 15.0 mL methanol was adjusted to 6 by the dropwise addition of acetic acid. Powered sodium cyanoborohydride (2.15 mmol) was then added, and the resultant mixture was refluxed for 5 h. After removal of the methanol by rotary evaporation, the residue was triturated in water and extracted with dichloromethane. The organic layer was washed with water, dried over magnesium sulfate and evaporated to dryness. Column chromatography of the residue on silica gel using ethyl acetate - cyclohexane (1/1) then methanol-chloroform (1/9) as eluents gave the crude product. This solid was then triturated with diethyl ether, filtered, washed with diethyl ether and dried under reduced pressure to give the compounds **1a-l**.

1.5.2.10.1. *Ethyl 4-{4-[(4-(2-oxo-2,3-dihydro-1H-benzimidazol-1-yl)piperidin-1-yl)benzyl]}-2-phenylpyrrolo[1,2-a]quinoxaline-3-carboxylate (**1a**)*. White crystals (53%), mp 186°C. ^1H NMR ($CDCl_3$) δ: 10.08 (s, 1H, NH), 8.11 (d, *J*=7.8 Hz, 1H, H-9), 8.09 (s, 1H, H-1), 7.97 (d, *J*=7.6 Hz, 1H, H-6), 7.79 (d, *J*=7.6 Hz, 2H, H-2' and H-6'), 7.64-7.54 (m, 4H, H-7, H-8, H-3' and H-5'), 7.46-7.37 (m, 5H, H-4 benzimid., H-2 phenyl, H-3 phenyl, H-5 phenyl and H-6 phenyl), 7.12-7.05 (m, 4H, H-4 phenyl, H-5 benzimid., H-6 benzimid. and H-7 benzimid.), 4.46-4.43 (m, 1H, CH pip.), 3.68 (q, *J*=6.9 Hz, 2H, CH_2), 3.50 (s, 2H, CH_2N), 3.14-3.09 (m, 2H, NCH_2 pip.), 2.53-2.24 (m, 4H, CH_2 pip. and NCH_2 pip.), 1.88-1.84 (m, 2H, CH_2 pip.), 0.94 (t, *J*=6.9 Hz, 3H, CH_3). ^{13}C NMR ($CDCl_3$) δ: 165.52 (C=O), 157.82 (C-4), 154.95 (C=O benzimid.), 140.15 (C-4'), 138.42 (C-1 phenyl, C-1' and C-5a), 136.42 (C-3a), 133.31 (C-6 and C-3a benzimid.), 130.42 (C-2', C-6' and C-7a benzimid.), 129.81 (C-7), 129.26 (C-3' and C-5'), 128.95 (C-8), 128.45 (C-3 phenyl, C-4 phenyl and C-5 phenyl), 127.99 (C-2 phenyl and C-6 phenyl), 127.64 (C-9a), 125.81 (C-2), 123.69 (C-3, C-5 benzimid. and C-6 benzimid.), 121.10 (C-1), 113.76 (C-9), 109.66 (C-4 benzimid. and C-7 benzimid.), 64.88 (CH_2N), 61.04 (CH_2), 53.15 (NCH_2 pip.), 50.71 (CH pip.), 29.23 (CH_2 pip.), 13.68 (CH_3). HRMS-ESI m/z $[M+Na]^+$ Calcd for $C_{39}H_{35}N_5O_3Na$: 644.2637, Found: 644.2634.

1.5.2.10.2. *Ethyl 4-{4-[(4-(5-fluoro-1H-benzimidazol-2-yl)piperidin-1-yl)benzyl]}-2-phenylpyrrolo[1,2-a]quinoxaline-3-carboxylate (**1b**)*. Cream crystals (61%), mp 163°C. ^{1}H NMR ($CDCl_3$) δ: 8.08 (s, 1H, H-1), 8.04 (d, *J*=7.8 Hz, 1H, H-9), 7.96 (d, *J*=7.8 Hz, 1H, H-6), 7.70 (d, *J*=7.2 Hz, 2H, H-2' and H-6'), 7.59 (t, *J*=7.2 Hz, 1H, H-7), 7.55-7.47 (m, 3H, H-3', H-5' and H-8), 7.44-7.33 (m, 7H, H-2 phenyl, H-3 phenyl, H-5 phenyl, H-6 phenyl, H-4 benzimid., H-6 benzimid. and H-7 benzimid.), 6.93 (t, *J*=7.2 Hz, 1H, H-4 phenyl), 3.61 (q, *J*=7.0 Hz, 2H, CH_2), 3.54 (s, 2H, CH_2N), 2.91 (d, 2H, *J*=10.2 Hz, NCH_2 pip.), 2.84 (t, *J*=11.4 Hz, 1H, CH pip.), 2.05 (t, *J*=10.8 Hz, 2H, CH_2 pip.), 1.99 (d, *J*=11.4 Hz, 2H, NCH_2 pip.), 1.94-1.85 (m, 2H, CH_2 pip.), 0.87 (t, *J*=7.0 Hz, 3H, CH_3). ^{13}C NMR ($CDCl_3$) δ: 165.68 (C=O), 160.05 (d, J_{CF}=234 Hz, C-5 benzimid.), 158.46 (C-4), 139.78 (C-4'), 138.14 (C-1 phenyl, C-1', C-2 benzimid., C-5a and C-3a benzimid.), 135.91 (C-3a and C-5a), 133.21 (C-6 and C-7a benzimid.), 130.55 (C-2' and C-6'), 129.99 (C-3' and C-5'), 129.23 (C-7), 128.91 (C-8), 128.51 (C-4 phenyl), 127.95 (C-2 phenyl, C-3 phenyl, C-5 phenyl and C-6 phenyl), 127.73 (C-9a), 125.85 (C-2), 123.62 (C-3), 113.95 (C-9), 113.48 (C-1 and C-7 benzimid.), 113.24 (C-4 benzimid. and C-6 benzimid.), 64.88 (CH_2N), 61.04 (CH_2), 53.15 (NCH_2 pip.), 50.71 (CH pip.), 29.23 (CH_2 pip.), 13.68 (CH_3). HRMS-ESI m/z $[M+H]^+$ Calcd for $C_{39}H_{35}N_5O_2F$: 624.2775, Found: 624.2751.

1.5.2.10.3. *Ethyl 4-{4-[4-(3-(pyridin-2-yl)-1,2,4-triazol-5-yl)piperidin-1-yl)benzyl]}-2-phenylpyrrolo[1,2-a]quinoxaline-3-carboxylate (**1c**)*. White crystals (56%), mp 143°C. ^{1}H NMR ($CDCl_3$) δ: 8.73 (d, *J*=3.0 Hz, 1H, H-6 pyr.), 8.22 (d, *J*=7.8 Hz, 1H, H-5 pyr.), 8.08 (d, *J*=8.4 Hz, 1H, H-6), 8.06 (s, 1H, H-1), 7.94 (d, *J*=8.4 Hz, 1H, H-9), 7.84 (t, *J*=7.8 Hz, 1H, H-4 pyr.), 7.74 (d, *J*=7.8 Hz, 2H, H-2' and H-6'), 7.58 (t, *J*=7.8 Hz, 1H, H-7), 7.56-7.49 (m, 4H, H-3', H-5', H-2 phenyl and H-6 phenyl), 7.44-7.40 (m, 2H, H-3 phenyl and H-5 phenyl), 7.38-7.34 (m, 2H, H-4 phenyl and H-3 pyr.), 7.28 (m, 1H, H-8), 3.65 (s, 2H, CH_2N), 3.64 (q, *J*=7.2 Hz, 2H, CH_2), 3.06 (d, 2H, *J*=10.8 Hz, NCH_2 pip.), 2.97-2.89 (m, 1H, CH pip.), 2.23 (t, *J*=10.2 Hz, 2H, CH_2 pip.), 2.12 (d, *J*=12.0 Hz, 2H, NCH_2 pip.), 2.10-2.00 (m, 2H, CH_2 pip.), 0.91 (t, *J*=7.2 Hz, 3H, CH_3). ^{13}C NMR ($CDCl_3$) δ: 165.56 (C=O), 154.36 (C-4, C-3 Triazole and C-2 pyr.), 149.50 (C-6 pyr.), 147.15 (C-5 Triazole), 138.36 (C-4' and C-4 pyr.), 137.48 (C-1 phenyl, C-1' and C-5a), 136.04 (C-3a), 133.31 (C-6), 130.40 (C-2' and C-6'), 129.41 (C-3', C-5' and C-7), 128.97 (C-8), 128.47 (C-4 phenyl), 128.31 (C-2 phenyl and C-6 phenyl), 127.95 (C-3 phenyl and C-5 phenyl), 127.62 (C-9a), 125.83 (C-2), 124.60 (C-3 pyr.), 123.69 (C-3), 121.79 (C-5 pyr.), 113.78 (C-9), 113.32 (C-1), 63.02 (CH_2N), 61.19 (CH_2), 53.39 (CH pip. and NCH_2), 30.74 (CH_2 pip.), 13.68 (CH_3). HRMS-ESI m/z $[M+Na]^+$ Calcd for $C_{39}H_{35}N_7O_2Na$: 656.2750, Found: 656.2768.

1.5.2.10.4. *Ethyl 4-{4-[(4-(2-oxo-2,3-dihydro-1H-benzimidazol-1-yl)piperidin-1-yl)benzyl]}-3-phenylpyrrolo[1,2-a]quinoxaline-2-carboxylate (**1d**)*. White crystals (60%), mp 260°C. ^{1}H NMR ($CDCl_3$) δ: 9.95 (s, 1H,NH), 8.67 (s, 1H, H-1), 8.07 (d, *J*=8.0 Hz, 1H, H-6), 8.03 (d, *J*=8.0 Hz, 1H, H-9), 7.61 (t, *J*=8.0 Hz, 1H, H-7), 7.55 (t, *J*=8.0 Hz, 1H, H-8), 7.34 (d, *J*=7.2 Hz, 1H, H-4 benzimid.), 7.23 (d, *J*=7.8 Hz, 2H, H-2' and H-6'), 7.15-7.07 (m, 3H, H-5 benzimid., H-6 benzimid., H-7 benzimid.), 7.06-6.96 (m, 7H, H-3', C-5', H-2 phenyl, H-3 phenyl, H-4 phenyl, H-5 phenyl and H-6 phenyl), 4.45 (t, 1H, CH pip.), 4.23 (q, 2H, CH_2), 3.46 (s, 2H, CH_2N), 3.00 (d, *J*=10.2 Hz, 2H, NCH_2 pip.), 2.52 (q, *J*=7.2 Hz, 2H, CH_2 pip.), 2.15 (t, *J*=11.4 Hz, 2H, NCH_2 pip.), 1.88 (d, *J*=10.8 Hz, 2H, CH_2 pip.), 1.19 (t, *J*=7.2 Hz, 3H, CH_3). ^{13}C NMR ($CDCl_3$) δ: 165.54 (C=O), 158.05 (C-4), 156.58 (C=O benzimid.), 140.33 (C-4'), 137.93 (C-1'), 137.56 (C-3a and C-5a), 135.22 (C-1 phenyl), 132.62 (C-3a benzimid.,

C-3 phenyl, C-4 phenyl and C-5 phenyl), 131.80 (C-6), 130.70 (C-7a benzimid.), 130.06 (2' and C-6'), 129.54 (C-7, C-3' and C-5'), 128.20 (C-2 phenyl and C-6 phenyl), 127.80 (C-8), 126.81 (C-9a), 124.41 (C-2), 122.61 (C-5 benzimid.), 122.52 (C-6 benzimid.), 120.66 (C-3), 119.65 (C-1), 115.21 (C-9) 111.25 (C-7 benzimid.), 111.17 (C-4 benzimid.), 64.03 (CH_2N), 61.79 (CH_2), 54.50 (NCH_2 pip.), 52.37 (CH pip.), 30.81 (CH_2 pip.), 15.53 (CH_3). HRMS-ESI m/z $[M+H]^+$ Calcd for $C_{39}H_{36}N_5O_3$: 622.2818, Found: 622.2795.

1.5.2.10.5. *Ethyl 4-{4-[(4-(5-fluoro-1H-benzimidazol-2-yl)piperidin-1-yl)benzyl]}-3-phenylpyrrolo[1,2-a]quinoxaline-2-carboxylate (**1e**)*. White crystals (78%), mp 246°C. ^{1}H NMR ($CDCl_3$) δ: 9.87 (s, 1H,NH), 8.67 (s, 1H, H-1), 8.04 (d, *J*=7.0 Hz, 1H, H-9), 8.03 (t, *J*=7.0 Hz, 1H, H-6), 7.61 (t, *J*=7.0 Hz, 1H, H-7), 7.54 (t, *J*=7.0 Hz, 1H, H-8), 7.18 (d, *J*=8.0 Hz, 2H, H-2' and H-6'), 7.06-6.96 (m, 7H, H-3', H-5', H-2 phenyl, H-3 phenyl, H-4 phenyl, H-5 phenyl, H-6 phenyl, H-4 benzimid., H-6 benzimid. and H-7 benzimid.), 4.23 (q, *J*=7.2 Hz, 2H, CH_2), 3.41 (s, 2H, CH_2N), 2.97-2.86 (m, 3H, NCH_2 pip. and CH pip.), 2.11 (d, *J*=11.6 Hz, 2H, CH_2 pip.), 2.08 (t, *J*=11.6 Hz, 2H, NCH_2 pip.), 1.93 (m, 2H, CH_2 pip.), 1.20 (t, *J*=7.2 Hz, 3H, CH_3). ^{13}C NMR ($CDCl_3$) δ: 165.56 (C=O), 161.10 (d, J_{CF}=236 Hz, C-5 benzimid.), 158.21 (C-4), 140.16 (C-4'), 138.00 (C-1 phenyl, C-1', C-2 benzimid. and C-3a benzimid.), 137.57 (C-3a and C-5a), 135.19 (C-7a benzimid.), 131.62 (C-6), 130.11 (C-2' and C-6'), 129.64 (3' and C-5'), 129.54 (C-7), 128.19 (C-8), 128.04 (C-2 phenyl, C-3 phenyl, C-4 phenyl, C-5 phenyl and C-6 phenyl), 127.77 (9a), 124.44 (C-2), 120.69 (C-3), 119.69 (C-1), 115.31 (C-9 and C-7 benzimid.), 111.90 (C-4 benzimid. and C-6 benzimid.), 64.24 (CH_2N), 61.92 (CH_2), 54.66 (NCH_2 pip.), 38.34 (CH pip.), 32.34 (CH_2 pip.), 15.60 (CH_3). HRMS-ESI m/z $[M+H]^+$ Calcd for $C_{39}H_{35}N_5O_2F$: 624.2775, Found: 624.2783.

1.5.2.10.6. *Ethyl 4-{4-[4-(3-(pyridin-2-yl)-1,2,4-triazol-5-yl)piperidin-1-yl)benzyl]}-3-phenylpyrrolo[1,2-a]quinoxaline-2-carboxylate (**1f**)*. Pale yellow crystals (53%), mp 139°C. ^{1}H NMR ($CDCl_3$) δ: 8.73 (d, *J*=3.0 Hz, 1H, H-6 pyr.), 8.64 (s, 1H, H-1), 8.23 (d, *J*=7.8 Hz, 1H, H-5 pyr.), 8.02 (d, *J*=8.4 Hz, 1H, H-9), 7.99 (d, *J*=8.4 Hz, 1H, H-6), 7.84 (t, *J*=7.8 Hz, 1H, H-4 pyr.), 7.56 (t, *J*=7.8 Hz, 1H, H-7), 7.50 (t, *J*=7.2 Hz, 1H, H-8), 7.36 (t, *J*=6.0 Hz, 1H, H-3 pyr.), 7.17 (d, *J*=7.2 Hz, 2H, H-2' and H-6'), 7.06-7.02 (m, 1H, H-4 phenyl),7.01-6.94 (m, 6H, H-3', H-5', H-2 phenyl, H-3 phenyl, H-5 phenyl, and H-6 phenyl), 4.20 (q, *J*=7.2 Hz, 2H, CH_2), 3.49 (s, 2H, CH_2N), 3.01-2.93 (m, 3H, NCH_2 pip. and CH pip.), 2.23-2.13 (m, 4H, NCH_2 pip. and CH_2 pip.), 2.12-2.03 (m, 2H, CH_2 pip.), 1.17 (t, *J*=7.2 Hz, 3H, CH_3). ^{13}C NMR ($CDCl_3$) δ: 164.07 (C=O), 156.45 (C-4, C-3 Triazole and C-2 pyr.), 149.47 (C-6 pyr.), 147.36 (C-5 Triazole), 138.47 (C-4' and C-4 pyr.), 136.63 (C-1 phenyl and C-1'), 135.97 (C-5a), 133.63 (C-3a), 130.21 (C-6), 128.64 (C-2' and C-6'), 128.48 (C-3' and C-5'), 128.09 (C-7), 126.77 (C-8), 126.49 (C-3 phenyl, C-4 phenyl and C-5 phenyl), 126.42 (C-2 phenyl and C-6 phenyl), 126.31 (9a), 125.41 (C-3 pyr.), 122.87 (C-2), 121.87 (C-5 pyr.), 119.12 (C-3), 118.26 (C-9), 113.76 (C-1), 62.64 (CH_2N), 60.34 (CH_2), 53.03 (NCH_2 pip.), 35.19 (CH pip.), 30.41 (CH_2 pip.), 14.07 (CH_3). HRMS-ESI m/z $[M+H]^+$ Calcd for $C_{39}H_{36}N_7O_2$: 634.2930, Found: 634.2952.

1.5.2.10.7. *Ethyl 4-{4-[(4-(2-oxo-2,3-dihydro-1H-benzimidazol-1-yl)piperidin-1-yl)benzyl]}-1-phenylpyrrolo[1,2-a]quinoxaline-3-carboxylate (**1g**)*. Pale yellow crystals (66%), mp 168°C.^{1}H NMR ($CDCl_3$) δ: 10.52 (s, 1H,NH), 8.10 (d, *J*=8.0 Hz, 1H, H-6), 7.82 (d, *J*=7.6 Hz, 2H, H-2' and H-6'), 7.62-7.53 (m, 8H, H-4 benzimid., H-2 phenyl, H-3 phenyl, H-4 phenyl, H-5 phenyl, H-6 phenyl, H-3' and H-5'), 7.44 (d, *J*=8.0 Hz,

1H, H-9), 7.43 (t, *J*=8.0 Hz, 1H, H-7), 7.33 (d, *J*=8.0 Hz, 1H, H-5 benzimid.), 7.21-7.02 (m, 4H, H-2, H-8, H-6 benzimid. and C-7 benzimid.), 4.52-4.36 (m, 1H, CH pip.), 3.79 (q, *J*=7.2 Hz, 2H, CH_2), 3.69 (s, 2H, CH_2N), 3.13 (d, *J*=10.8 Hz, 2H, NCH_2 pip.), 2.62-2.48 (m, 2H, CH_2 pip.), 2.27 (t, *J*=10.8 Hz, 2H, NCH_2 pip.), 1.86 (d, *J*=14.8 Hz, 2H, CH_2 pip.), 0.98 (t, *J*=7.2 Hz, 3H, CH_3). ^{13}C NMR ($CDCl_3$) δ: 166.54 (C=O), 156.74 (C=O benzimid.), 155.97 (C-4), 141.21 (C-3a and C-4'), 140.77 (C-5a), 134.78 (C-1'), 133.44 (C-1 phenyl and C-3a benzimid.), 131.68 (C-6), 131.12 (C-3' and C-5'), 130.56 (C-7a benzimid.), 130.37 (C-2 phenyl, C-3 phenyl, C-4 phenyl, C-5 phenyl and C-6 phenyl), 129.64 (C-2' and C-6'), 129.46 (C-7), 128.66 (C-8), 127.22 (C-9a), 122.56 (C-1, C-3 and C-5 benzimid.), 120.22 (C-2 and C-6 benzimid.), 118.10 (C-9), 111.22 (C-7 benzimid.), 111.19 (C-4 benzimid.), 64.09 (CH_2N), 62.27 (CH_2), 54.60 (NCH_2 pip.), 52.18 (CH pip.), 30.69 (CH_2 pip.), 15.27 (CH_3). HRMS-ESI m/z $[M+H]^+$ Calcd for $C_{39}H_{36}N_5O_3$: 622.2818, Found: 622.2802.

1.5.2.10.8. *Ethyl 4-{4-[(4-(5-fluoro-1H-benzimidazol-2-yl)piperidin-1-yl)benzyl]}-1-phenylpyrrolo[1,2-a]quinoxaline-3-carboxylate (**1h**)*. Pale yellow crystals (60%), mp 157°C. ^{1}H NMR ($CDCl_3$) δ: 11.40 (s, 1H,NH), 8.02 (d, *J*=7.6 Hz, 1H, H-6), 7.74 (d, *J*=8.0 Hz, 2H, H-2' and H-6'), 7.61-7.52 (m, 5H, H-2 phenyl, H-3 phenyl, H-4 phenyl, H-5 phenyl and H-6 phenyl), 7.46 (d, *J*=8.0 Hz, 1H, H-9), 7.44-7.37 (m, 3H, H-7, H-3' and H-5'), 7.21-7.15 (m, 4H, H-2, H-8, H-4 benzimid. and H-7 benzimid.), 6.96 (td, *J*=9.2 Hz and 1.6 Hz, 1H, H-6 benzimid.), 3.72 (q, *J*=7.2, 2H, CH_2), 3.55 (s, 2H, CH_2N), 2.98-2.84 (m, 3H, NCH_2 pip. and CH pip.), 2.13-1.98 (m, 4H, NCH_2 pip. and CH_2 pip.), 1.96-1.81 (m, 2H, CH_2 pip.), 0.92 (t, *J*=7.2 Hz, 3H, CH_3). ^{13}C NMR ($CDCl_3$) δ: 166.61 (C=O), 160.7 (d, J_{CF}=235 Hz, C-5 benzimid.), 156.16 (C-4), 141.49 (C-4'), 140.40 (C-3a and C-5a), 138.66 (C-2 benzimid. and C-3a benzimid.), 134.65 (C-1 phenyl), 133.67 (C-7a benzimid.), 131.22 (C-3' and C-5'), 131.11 (C-3 phenyl, C-4 phenyl and C-5 phenyl), 130.63 (C-6, C-2 phenyl and C-6 phenyl), 130.39 (C-7), 129.35 (C-2' and C-6'), 128.72 (C-8), 128.54 (C-9a), 127.38 (C-1, C-3 and C-7 benzimid.), 120.33 (C-6 benzimid.), 118.21 (C-9), 111.50 (C-4 benzimid.), 64.19 (CH_2N), 62.32 (CH_2), 54.70 (NCH_2 pip.), 38.34 (CH pip.), 32.34 (CH_2 pip.), 15.21 (CH_3). HRMS-ESI m/z $[M+H]^+$ Calcd for $C_{39}H_{35}N_5O_2F$: 624.2775, Found: 624.2800.

1.5.2.10.9. *Ethyl 4-{4-[4-(3-(pyridin-2-yl)-1,2,4-triazol-5-yl)piperidin-1-yl)benzyl]}-1-phenylpyrrolo[1,2-a]quinoxaline-3-carboxylate (**1i**)*.Pale yellow crystals (69%), mp 146°C. ^{1}H NMR ($CDCl_3$) δ: 9.12 (s, 1H, NH), 8.66 (d, *J*=4.0 Hz, 1H, H-6 pyr.), 8.19 (d, *J*=8.0 Hz, 1H, H-5 pyr.), 8.07 (dd, *J*=8.2 and 1.6 Hz, 1H, H-6), 7.86 (td, *J*=7.6 and 1.6 Hz, 1H, H-4 pyr.), 7.79 (d, *J*=7.4 Hz, 2H, H-2' and H-6'), 7.58-7.55 (m, 5H, H-2 phenyl, H-3 phenyl, H-4 phenyl, H-5 phenyl and H-6 phenyl), 7.53 (d, *J*=7.4 Hz, 2H, H-3' and H-5'), 7.45 (d, *J*=8.2 Hz, 1H, H-9), 7.43 (t, *J*=7.2 Hz, 1H, H-7), 7.40 (dd, *J*=6.4 and 1.2 Hz, 1H, H-9), 7.37 (d, *J*=7.2 Hz, 1H, H-3 pyr.), 7.18 (td, *J*=7.2 Hz and 1.6 Hz, 1H, H-8), 7.17 (sl, 1H, H-2), 3.78 (q, *J*=7.2, 2H, CH_2), 3.71 (s, 2H, CH_2N), 3.10 (d, *J*=11.2 Hz, 2H, NCH_2 pip.), 3.05-2.88 (m, 1H, CH pip.), 2.19 (t, *J*=8.0 Hz, 2H, CH_2 pip.), 0.97 (t, *J*=7.2 Hz, 3H, CH_3). ^{13}C NMR ($CDCl_3$) δ: 156.67 (C-4 and C-2 pyr.), 152.37 (C-3 Triazole), 151.01 (C-6 pyr.), 147.10 (C-5 Triazole), 142.04 (C-1'), 141.03 (C-4'), 138.80 (C-4 pyr.), 134.91 (C-1 phenyl), 133.39 (C-5a), 131.80 (C-6), 130.85 (C-3' and C-5'), 130.59 (C-2 phenyl, C-3 phenyl, C-4 phenyl, C-5 phenyl and C-6 phenyl), 129.91 (C-9a, C-2' and C-6'), 128.43 (C-8), 127.27 (C-7 and C-3 pyr.), 126.22 (C-1 and C-3), 122.91 (C-5 pyr.), 120.27 (C-2), 118.17 (C-9), 62.30 (CH_2N), 61.18 (CH_2), 56.03 (CH pip.), 55.73 (NCH_2 pip.), 32.27 (CH_2 pip.), 15.36 (CH_3). HRMS-ESI m/z $[M+H]^+$ Calcd for $C_{39}H_{36}N_7O_2$: 634.2930, Found: 634.2959.

1.5.2.10.10. *Ethyl 4-{4-[(4-(2-oxo-2,3-dihydro-1H-benzimidazol-1-yl)piperidin-1-yl)benzyl]}-1-phenylpyrrolo[1,2-a]quinoxaline-2-carboxylate (**1j**)*. White crystals (70%), mp 166°C. ^{1}H NMR ($CDCl_3$) δ: 10.19 (s, 1H, NH), 8.03 (d, *J*=7.8 Hz, 1H, H-6), 8.02 (d, *J*=7.8 Hz, 2H, H-2' and H-6'), 7.67-7.57 (m, 6H, H-7, H-3 phenyl, H-4 phenyl, H-5 phenyl, H-3' and H-5'), 7.55 (s, 1H, H-3), 7.54-7.51 (m, 2H, H-2 phenyl and H-6 phenyl), 7.41-7.33 (m, 2H, H-9 and H-4 benzimid.), 7.14-7.05 (m, 4H, H-8, H-5 benzimid., H-6 benzimid., H-7 benzimid.), 4.47 (t, *J*=12.0 Hz, 1H, CH pip.), 4.17 (q, *J*=7.2 Hz, 2H, CH_2), 3.73 (s, 2H, CH_2N), 3.23-3.08 (m, 2H, NCH_2 pip.), 2.65-2.50 (m, 2H, CH_2 pip.), 2.37-2.23 (m, 2H, NCH_2 pip.), 1.89 (d, *J*=10.2 Hz, 2H, CH_2 pip.), 1.13 (t, *J*=7.2 Hz, 3H, CH_3).^{13}C NMR ($CDCl_3$) δ: 165.59 (C=O), 156.75 (C-4), 156.60 (C=O benzimid.), 139.29 (C-5a and C-3a benzimid.), 138.22 (C-1'), 136.20 (C-3a and C-1 phenyl), 131.88 (C-6), 131.85 (C-2 phenyl and C-6 phenyl), 130.85 (C-7a benzimid.), 130.80 (C-4 phenyl), 130.51 (C-3' and C-5'), 130.24 (C-7), 130.12 (C-3 phenyl and C-5 phenyl), 129.67 (C-2' and C-6'), 128.19 (C-8), 127.34 (C-9), 126.61 (C-2), 120.81 (C-1), 118.12 (C-5 benzimid. and C-6 benzimid.), 111.94 (C-3 and C-4 benzimid.), 111.20 (C-7 benzimid.), 64.01 (CH_2N), 61.61 (CH_2), 54.61 (NCH_2 pip.), 52.15 (CH pip.), 30.67 (CH_2 pip.), 15.42 (CH_3). HRMS-ESI m/z $[M+H]^+$ Calcd for $C_{39}H_{36}N_5O_3$: 622.2818, Found: 622.2846.

1.5.2.10.11. *Ethyl 4-{4-[(4-(5-fluoro-1H-benzimidazol-2-yl)piperidin-1-yl)benzyl]}-1-phenylpyrrolo[1,2-a]quinoxaline-2-carboxylate (**1k**)*. White crystals (46%), mp 196°C. ^{1}H NMR ($CDCl_3$) δ: 7.97 (d, *J*=7.8 Hz, 1H, H-6), 7.93 (d, *J*=7.2 Hz, 2H, H-2' and H-6'), 7.64-7.56 (m, 3H, H-7, H-9 and H-4 phenyl), 7.55-7.47 (m, 6H, H-3, H-3 phenyl, H-5 phenyl, H-3', H-5' and H-4 benzimid.), 7.36 (t, *J*=6.6 Hz, 2H, H-2 phenyl and H-6 phenyl), 7.09 (d, *J*=7.8 Hz, 1H, H-8), 7.08 (d, *J*=7.8 Hz, 1H, H-7 benzimid.), 6.98 (t, *J*=7.8 Hz, 1H, H-6 benzimid.), 4.16 (q, *J*=7.2 Hz, 2H, CH_2), 3.62 (s, 2H, CH_2N), 2.99 (d, *J*=11.4 Hz, 2H, NCH_2 pip.), 2.94 (t, *J*=12 Hz, 1H, CH pip.), 2.15 (t, *J*=11.4 Hz, 2H, NCH_2 pip.), 2.06 (d, *J*=11.4 Hz, 2H, CH_2 pip.), 1.98-1.89 (m, 2H, CH_2 pip.), 1.11 (t, *J*=7.2 Hz, 3H, CH_3). ^{13}C NMR ($CDCl_3$) δ: 165.75 (C=O), 160.57 (d, J_{CF}=235 Hz, C-5 benzimid.), 156.90 (C-4), 142.10 (C-4'), 139.21 (C-5a), 137.97 (C-1'), 136.42 (C-2 benzimid. and C-3a benzimid.), 134.60 (C-1 phenyl), 132.51(C-9a), 131.90 (C-3' and C-5'), 131.52 (C-3a), 130.91 (C-6), 130.84 (C-3 phenyl and C-5 phenyl), 130.32 (C-7a benzimid.), 130.05 (C-7, C-9 and C-4 phenyl.), 129.58 (C-2' and C-6'), 128.37 (C-8), 127.52 (C-2 phenyl and C-6 phenyl), 120.93 (C-1 and C-2), 118.28 (C-7 benzimid.), 111.82 (C-3 and C-6 benzimid.), 111.64 (C-4 benzimid.), 64.24 (CH_2N), 61.79 (CH_2), 54.74 (NCH_2 pip.), 38.31 (CH pip.), 32.35 (CH_2 pip.), 15.41 (CH_3). HRMS-ESI m/z $[M+H]^+$ Calcd for $C_{39}H_{35}N_5O_2F$: 624.2775, Found: 624.2747.

1.5.2.10.12. *Ethyl 4-{4-[4-(3-(pyridin-2-yl)-1,2,4-triazol-5-yl)piperidin-1-yl)benzyl]}-1-phenylpyrrolo[1,2-a]quinoxaline-2-carboxylate (**1l**)*. White crystals (42%), mp 145°C. ^{1}H NMR ($CDCl_3$) δ: 8.69 (d, *J*=4.6 Hz, 1H, H-6 pyr.), 8.19 (d, *J*=7.8 Hz, 1H, H-5 pyr.), 7.99 (d, *J*=7.0 Hz, 1H, H-6), 7.94 (d, *J*=7.0 Hz, 2H, H-2' and H-6'), 7.83 (td, *J*=7.6 and 1.6 Hz, 1H, H-4 pyr.), 7.62-7.45 (m, 8H, H-3, H-7, H-9, H-3 phenyl, H-4 phenyl, H-5 phenyl, H-3' and H-5'), 7.40-7.29 (m, 3H, H-2 phenyl, H-6 phenyl and H-3 pyr.), 7.04 (d, *J*=7.0 Hz, 1H, H-8), 4.14 (q, *J*=7.1 Hz, 2H, CH_2), 3.67 (s, 2H, CH_2N), 3.07 (d, *J*=11.0 Hz, 2H, NCH_2 pip.), 2.99-2.79 (m, 1H, CH pip.), 2.34-1.90 (m, 6H, NCH_2 pip. and 2 x CH_2 pip.), 1.10 (t, *J*=7.1 Hz, 3H, CH_3). ^{13}C NMR ($CDCl_3$) δ: 165.65 (C=O), 156.79 (C-4 and C-2 pyr.), 150.99 (C-3 Triazole and C-6 pyr.), 148.41 (C-5 Triazole), 139.39 (C-4'), 138.86 (C-5a and C-4 pyr.), 138.18 (C-1'), 134.72 (C-1 phenyl), 131.94

(C-3a, C-3' and C-5'), 130.98 (C-6), 130.86 (C-3 phenyl and C-5 phenyl), 130.11 (C-7, C-9 and C-4 phenyl), 129.56 (C-2' and C-6'), 128.22 (C-8), 127.37 (C-3 pyr., C-2 phenyl and C-6 phenyl), 123.11 (C-2, C-9a and C-5 pyr.), 120.84 (C-1), 112.01 (C-3), 64.51 (CH_2N), 61.68 (CH_2), 55.03 (NCH_2 pip.), 32.32 (CH_2 pip. and CH pip.), 15.49 (CH_3). HRMS-ESI m/z $[M+H]^+$ Calcd for $C_{39}H_{36}N_7O_2$: 634.2930, Found: 634.2962.

1.5.3. Melting point, IR, TLC, LC, and ESI-MS

Melting points were determined with an SM-LUX-POL Leitz hot-stage microscope and are uncorrected. IR spectra were recorded on a NICOLET 380FT-IR spectrophotometer. Analytical TLC were carried out on 0.25 precoated silica gel plates (POLYGRAM SIL G/UV_{254}) and visualization of compounds after UV light irradiation. Silica gel 60 (70-230 mesh) was used for column chromatography. High resolution mass spectra (electrospray in positive mode, ESI+) were recorded on a Waters Q-TOF Ultima apparatus. Elemental analyses were found within ±0.4% of the theoretical values.

1.5.4. NMR experiments

All the synthetic molecules were characterized by 1D [^{1}H] and 2D [^{1}H, ^{1}H] NMR spectroscopy. Different types of NMR experiments were recorded to obtain a full assignment of ^{1}H and ^{13}C atoms. Most ^{1}H and ^{13}C NMR spectra were recorded at 300 MHz on a Bruker Avance 300 spectrometer equipped with a z-gradient 5 mm triple-resonance probe. Additional proton spectra were recorded on a 600 MHz Varian instrument provided with a cold-probe. Experiments were acquired at 298 K; samples consisted of compounds (~2 mg) dissolved in 600 µl of Chloroform-d (99.8% D, Sigma-Aldrich, USA). 1D proton spectra were recorded with 64-128 scans and a relaxation delay of 1 s. 1D, proton decoupled, ^{13}C spectra were acquired with 7000-10000 scans and d1=5-10 s. 2D [^{1}H, ^{1}H] TOCSY (Total Correlation Spectroscopy) [38] and 2D [^{1}H, ^{1}H] ROESY (Rotating frame Overhauser Enhancement Spectroscopy) [39] experiments (2048*256 total data points and 32-64 scans per t1 increment) were recorded with mixing times of 70 ms and 200 ms respectively. A 2D [^{1}H, ^{1}H] DQFCOSY (2048*256 total data points and 64-128 scans pet t1 increment) (Double Quantum Filter Correlation Spectroscopy) [40] experiment was also acquired to complete proton chemical shift assignment. [^{1}H, ^{13}C] HMBC [41] (1024*128-256 total data points and 128 scans pet t1 increment) and [^{1}H, ^{13}C] HMQC [42] (1024*128 total data points and 64-128 scans pet t1 increment) spectra were recorded to complete and verify ^{13}C assignments. Chemical shifts were referenced to residual $CDCl_3$ at 7.24 ppm for ^{1}H spectra and 77.0 ppm for ^{13}C experiments. Spectra were processed with Varian software (vnmrj_1.1D). 2D experiments were analyzed with the program NEASY [43] as implemented in Cara (http://www.nmr.ch). Chemical shift values are compatible with those previously reported for similar molecules. [14] Splitting patterns have been designated as follows: s = singlet; bs = broad singlet; d = doublet; t = triplet; q = quartet; dd = double doublet; ddd = double double doublet; dt = double triplet; m = multiplet.

1.5.5. X-ray Data

Colorless single crystal of **2b** was obtained by slow evaporation from chloroform: monoclinic, space group $P2_1/c$, a=11.9390(11) Å, b=13.8893(9) Å, c=18.5001(19) Å, α=90°, β=100.959(7)°, γ=90°, V=3011.8(5) Å^3, Z=8, δ(calcd)= 1.267 Mg.m^{-3}, FW=287.31 for $C_{16}H_{17}NO_4$, F(000)=1216. Colorless single crystal of **2c** was obtained by slow evaporation from methanol/dichloromethane (20/80) solution: monoclinic,

space group $P2_1/n$, *a*=14.653(3) Å, *b*=5.1458(12) Å, *c*=18.842(4) Å, α=90°, β=94.447(14)°, γ=90°, *V*=1416.4(5) Å^3, Z=4, δ(calcd)= 1.347 Mg.m^{-3}, FW=287.31 for $C_{16}H_{17}NO_4$, *F*(000)=608. Colorless single crystal of **2d** was obtained by slow evaporation from methanol/dichloromethane (20/80) solution: triclinic, space group $P\text{-}1$, *a*=8.628(9) Å, *b*=9.798(2) Å, *c*=10.095(3) Å, α=73.02(2)°, β=71.00(3)°, γ=73.95(5)°, *V*=756.1(8) Å^3, Z=1, δ(calcd)= 1.262 Mg.m^{-3}, FW=574.61 for $C_{32}H_{34}N_2O_8$, *F*(000)=304. Colorless single crystal of **1a** was obtained by slow evaporation from methanol/dichloromethane (30/70) solution: triclinic, space group $P\text{-}1$, *a*=10.0140(10) Å, *b*=10.877(2) Å, *c*=17.580(2) Å, α=79.642(9)°, β=86.041(8)°, γ=65.016(8)°, *V*=1707.3(4) Å^3, Z=2, δ(calcd)= 1.311 Mg.m^{-3}, FW=673.96 for $C_{39}H_{35}N_5O_3$,0.48($CH_2Cl_2O_2$),0.19(O), *F*(000)=710. Colorless single crystal of **1d** was obtained by slow evaporation from methanol/dichloromethane (30/70) solution: triclinic, space group $P\text{-}1$, *a*=8.3759(10) Å, *b*=12.6418(13) Å, *c*=16.6718(18) Å, α=109.502(7)°, β=96.482(8)°, γ=103.928(7)°, *V*=1578.7(3) Å^3, Z=2, δ(calcd)= 1.308 Mg.m^{-3}, FW=621.72 for $C_{39}H_{35}N_5O_3$, *F*(000)=656. Colorless single crystal of **1g** was obtained by slow evaporation from methanol/dichloromethane (20/80) solution: triclinic, space group $P\text{-}1$, *a*=9.6882(14) Å, *b*=11.2689(13) Å, *c*=19.639(2) Å, α=77.504(8)°, β=83.067(11)°, γ=72.264(11)°, *V*=1990.4(4) Å^3, Z=2, δ(calcd)= 1.347 Mg.m^{-3}, FW=807.16 for $C_{39}H_{35}N_5O_3$,1.56($CHCl_3$), *F*(000)=836. Colorless single crystal of **1j** was obtained by slow evaporation from methanol/dichloromethane (30/70) solution: triclinic, space group $P\text{-}1$, *a*=9.7518(8) Å, *b*=11.1004(8) Å, *c*=17.9247(12) Å, α=86.622(5)°, β=82.713(5)°, γ=66.141(5)°, *V*=1760.1(2) Å^3, Z=2, δ(calcd)= 1.398 Mg.m^{-3}, FW=741.09 for $C_{39}H_{35}N_5O_3$,$CHCl_3$, *F*(000)=772. Pale-yellow single crystal of **1k** was obtained by slow evaporation from methanol/dichloromethane (30/70) solution: monoclinic, space group C2/c, *a*=18.0617(14) Å, *b*=9.4210(8) Å, *c*=43.021(3) Å, α=90°, β= 93.345(5)°, γ=90°, *V*=7307.9(10) Å^3, Z=8, δ(calcd)= 1.296 Mg.m^{-3}, FW=712.78 for $C_{39}H_{34}FN_5O_2$,5(H_2O) *F*(000)=3016. Full crystallographic results have been deposited at the Cambridge Crystallographic Data Centre (CCDC-1014944, CCDC-891817, CCDC-891816, CCDC-891812, CCDC-891811, CCDC-891813, CCDC-891814, CCDC-891815, respectively), UK, as supplementary material. [44] The data were corrected for Lorentz and polarization effects and for empirical absorption correction. [45] The structure was solved by direct methods Shelx 2013 [46] and refined using Shelx 2013 [46] suite of programs.

1.5.6. Biological activity

1.5.6.1. Cell culture

The human leukemic cell lines U937, K562, HL60, U266 and Jurkat were grown in RPMI 1640 medium (Life Technology, France) supplemented with 10% fetal calf serum (FCS), antibiotics (100 U/mL penicillin, 100 µg/mL streptomycin) and L-glutamin, (Eurobio, France) at 37°C, 5% CO_2 in air. The toxicity of various molecules was also evaluated on non-activated, freshly isolated normal human peripheral blood mononuclear cells (PBMNC), as well as phytohemagglutinin (T lymphoproliferative agent) (PHA)-induced cells. PBMNC from blood of healthy volunteers were obtained following centrifugation on Ficoll gradient. Cells were then incubated in medium alone or induced to enter cell cycle by the addition of PHA (5.00 µg/mL, Murex Biotech Limited, Dartford, UK).

1.5.6.2. Cytotoxicity Test

The MTS cell proliferation assay (Promega, France) is a colorimetric assay system, which measures the reduction of a tetrazolium component (MTS) into formazan produced by the mitochondria of viable cells. Cells were washed twice in PBS (Phosphate Buffer Saline) and plated in quadruplicate into microtiter-plate wells in 100 μL culture media with or without our various compounds at increasing concentrations (0, 1, 5, 10, 20 and 50 μM) during 1, 2 and 3 days. After 3 hours of incubation at 37°C with 20 μL MTS/well, the plates were read by using an ELISA microplate reader (Thermo, Electrocorporation) at 490 nm wavelength. The amount of colour produced was directly proportional to the number of viable cells. The results are expressed as the concentrations inhibiting cell growth by 50% after a 3 days incubation period. The 50% cytotoxic concentrations (CC_{50}) were determined by linear regression analysis, expressed in μM ± SD (Microsoft Excel).

1.5.7. Computational prediction of toxicity and drug relevant properties

Calculations of Clog *P* and calculations of Topological Polar Surface Area (TPSA), number of hydrogen bond acceptor (nON) and donor (nNH/OH) atoms, and any violations to the Lipinski's "rule of five" (logP ≤5, molecular weight ≤500, number of hydrogen bond acceptors ≤10, and number of hydrogen bond donors ≤5 [46]; were performed using the MIPC server at http://www.molinspiration.com/cgi-bin/properties. [47]

1.5.8. Docking Studies

Computational dockingstudies were conducted withthe Swiss Dock server (http://www.swissdock.ch) that is based on the software EADock DSS [17], the Evolutionary Algorithm for Docking. Evolutionary algorithms are iterative stochastic optimization procedures where an initial population of solutions is generated and evaluated with respect to a set of constraints, described by the fitness function. In the docking problem, the fitness function describes the interactions between the ligand and the receptor.

For preliminary docking studies the input target consisted of the entire human Akt1 protein (including PH and Kinase domains) extracted from the crystal structure of the complex Akt1/allosteric inhibitor (pdb code: 3O96 [5], (***Figure 8b***). All the molecules considered as ligands (Series D, ***Figure 6***) and used as input files in the SwissDock server were generated with Chem3DPro 11.0 and then converted in .mol2 format with UCSF Chimera [48].

Next, docking studies were carried out by considering only the Akt PH domain. The input target consideredwas the PH domainextracted from the crystal structure of the complex PKBα/IP4 (pdb code: 1UNQ [19]), where IP4 represents the skeleton of the phosphatidylinositol (3,4,5) triphosphate (PIP3), a natural ligand of the Akt PH domain. (***Figure 23***)

2. Workpackage 2: CXCR4-Intrinsically disordered peptide ligand system

2.1. Background

2.1.1. Chemokines

Chemokines are a superfamily of 8-12-kDa hemoattractive cytokines constitutively secreted by stromal cells, including fibroblasts and endothelial cells. At present, >50 chemokines have been identified and they can be divided into four groups (C, CC, CXC and CX3C) based on the number and position of conserved cysteines, where C represents the number of cysteine residues and X denotes the number of amino acids between the conserved cysteines. Chemokines were initially discovered as essential mediators in the process of the directional migration of leukocytes to the infection and inflammation sites and have been increasingly demonstrated to regulate tumor development and metastasis [49].

To date, at least 20 chemokine receptors are already reported and classified into four subtypes [CXC chemokine receptors (CXCRs), CC chemokine receptors (CCRs), XCR and CX3CR] on the basis of their specific preference for some chemokines. It is confirmed that these chemokines can bind to different receptors but also that more than one chemokine is able to bind to the same receptor to a certain extent [49]. However, certain chemokines only interact with a single receptor. The binding of chemokines to their receptors stimulates the activation of several downstream signaling pathways that regulate tumor progression and metastasis [49].

2.1.2. CXCR4

Chemokine receptor 4 (CXCR4) is a rhodopsin-like seven-transmembrane G-protein coupled receptor, initially discovered as co-receptor facilitating the entry of T-tropic (X4) HIV viruses into CD4+ T cells. It is exclusively binds to CXCL12 [50]. It has been found that CXCR4 is expressed in a wide range of tissues, including brain, lymph node and small intestine tissues [50], as well as in monocytes, B cells, naïve T cells and early hematopoietic progenitor cells in the immune system [49]. CXCR4 receptor activation is mediated by coupling to an intracellular heterotrimeric G-protein associated with the inner surface of the plasma membrane [51].

2.1.3. CXCL12

Chemokine 12 (CXCL12), also designated as stromal cell-derived factor-1 (SDF-1), as mentioned above is a member of the CXC subfamily of chemokines. It is widely expressed in a number of organs, such as the lungs, liver, skeletal muscle, brain, kidneys, heart, skin and bone marrow. The involvement of CXCL12 in the metastasis of various types of cancer has also been previously demonstrated. [52] However, it has been reported that there are minimal or negligible effects on the survival and growth of myeloma in the presence of CXCL12 *in vitro*. [53] On the one hand, the expression of CXCL12 can be affected by a number of factors such as DNA-damaging agents and granulocyte colony-stimulating factor (G-CSF) [54].

On the other hand, CXCL12 can stimulate the secretion of other factors. It has previously been demonstrated that matrix metalloproteinase-9 (MMP-9) expression is upregulated in the presence of CXCL12 when investigating the involvement of the CXCL12-CXCR4 axis in the metastasis of prostate cancer and OS [55].

2.1.4. Downstream pathways involved in the CXCL12-CXCR4/CXCR7 interaction

It is already accepted that the binding of CXCL12 to CXCR4 or CXCR7 leads to the activation of several downstream pathways that regulate cell chemotaxis, survival, proliferation and migration [49]. (**Figure 23**)

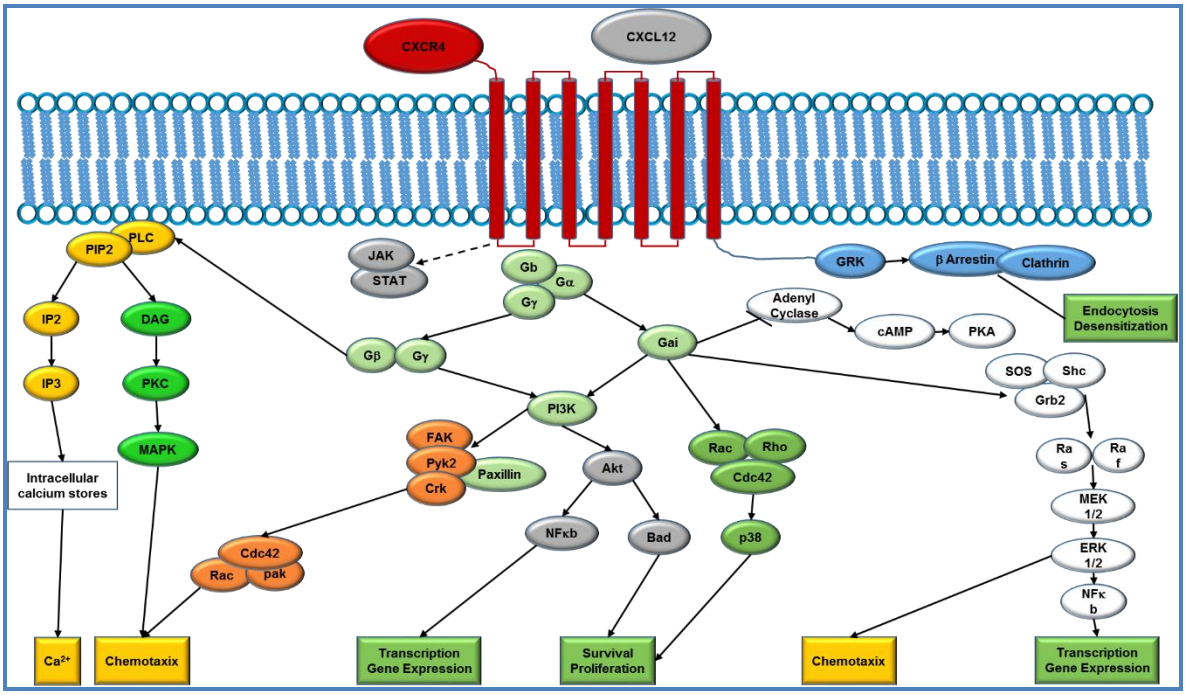


***Figure 23**. Scheme of the CXCL12/CXCR4 intracellular signal transduction pathways.*

Among them, it is useful to describe the mechanism involving adenylate cyclase (**see Figure 24a**). CXCR4 after the binding to CXCL12 promotes the conformational change of the Gα subunit and replacement of the bound GDP by guanine nucleotide triphosphate (GTP). This exchange triggers the further conformation changes within the Gα subunit, which allows the trimeric G protein to be released from the receptor, and to dissociate into the GTP-bound Gα subunit and Gβ/Gγ dimer. Both the activated components interact with various effector proteins. In particular, Gα regulates adenylyl cyclase (AC) that serves as an effector enzyme that catalyzes 5′adenosine triphosphate into cyclic adenosine monophosphate (cAMP) and thereby activates cAMP-dependent protein kinase, which regulates a host of other downstream effectors including MAPK signaling pathway. When a peptide antagonist binds to CXCR4, it inhibits this mechanism. (**Figure 24b**)

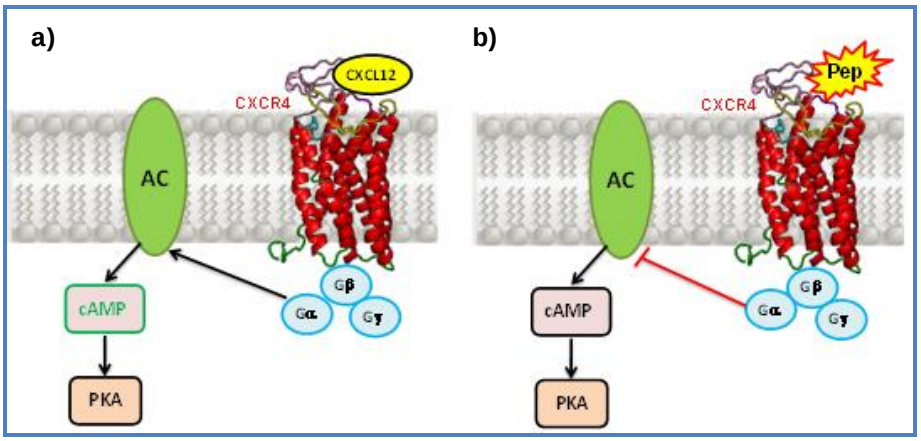


***Figure 24**. Scheme of the CXCL12/CXCR4 pathway with the adenylate cyclase **a)** when CXCL12 promote the receptor activity and **b)** when a peptide antagonist inhibits CXCR4.*

Because CXCR4 is overexpressed in several human cancers, the blockade of CXCR4–CXCL12 interactions has been extensively investigated as a potential cancer therapeutic. The lead CXCR4 antagonist, plerixafor (previously known as AMD3100), is the most clinically advanced compound. Since December 2008, it is an FDA-approved drug, in combination with granulocytecolony-stimulating factor, for hematopoietic stem and progenitorcell mobilization in patients with non-Hodgkin lymphoma and multiple myeloma even if it is a metal-chelating bicyclamthat has been shown to determine cardiotoxicity as reported inits clinical trial against HIV. Therefore, this suggests the need for new antagonists able to block, without any other compliance, the binding between CXCL12 and CXCR4 and/or that between CXCL12 and CXCR7 [51].

2.1.5. IDP

Taking into account the high specificity that the peptides could have [57] the development of therapeutic tools and diagnostic peptide-besed have recently gathered interest. In this cotext the intrinsically disorder peptides (IDP) currently represent an innovative and interesting starting point. More detailed analysis revealed that in comparison with “normal” ordered proteins, IDPs were significantly depleted in so-called order-promoting amino acids such as Ile, Leu, Val, Trp, Tyr, Phe, Cys, and Asn, being substantially enriched in disorder-promotingamino acids, Ala, Arg, Gly, Gln, Ser, Glu, Lys, and Pro [58]. Their peculiar flexibility results in the formation of transient complexes with binding partners characterized by low affinity because of the entropic cost associated with the disorder-to-order transition upon binding. A control of this transition between different structural states of an IDP may be useful in different fields. A possible way to achieve this goal consists in the peptide amphiphiles, a class of molecules in which a bioactive peptide is covalently conjugated to a hydrophobic moiety. This characteristic gives to PAs the ability to self-assemble spontaneously into a variety of nanostructures. In this way PAs could be considered as useful tools to overcome different cell barriers and to improve the ligand specificity against CXCR4 receptor. In a recent work, an IDP (R11) and its two lipophilic derivates PAs, $(C18)_2$-R11 and $(C18)_2$-L-R11and their structural properties in solution were investigated using CD and NMR [n= ref.]. As described in this study, $(C18)_2$-R11 and $(C18)_2$-L-R11display typical features of PAs, such as the formation of micelles and unilamellar vesicles. In addition, their surface properties were studied using Langmuir monomolecular films and the results obtained support the formation of molecular aggregates upon compression of the PA films. The presence of the alkyl chains induces not only the self-assembly of these new PAs into supramolecular aggregates but also a gain of structure within the disordered peptide [57].

2.2. Aim of the Workpackage II

The chemokines and their receptors play a key role in immune and inflammatory responses by promoting recruitment and activation of different sub populations of leukocytes. In this context, we focused our interest on CXCR4, a G-protein-coupled receptor involved in a number of physiological processes in the hematopoietic and immune systems. A recent work has described the structural behaviour and the biological binding between CXCR4 and its natural ligand CXCL12 [51]. This axis possesses a pivotal role in diseases, such as HIV, cancer, WHIM syndrome, rheumatoid arthritis, pulmonary fibrosis, and lupus and, hence, indicated as putative therapeutic target [61]. Nevertheless, the therapeutic utility of blocking the interaction of the chemokine receptors with their natural ligands is not so obvious. In effect, although they are markers of disease progression, it is not possible to say that they are its sole cause [59]. However, Horuk suggested that, to obtain an effective therapeutic role by inhibitory molecules, a potential solution could be to develop drugs that target more than one receptor [60]. Considering that CXCL12 binds to CXCR4 as well as CXCR7, we decided to design new linear peptides, charaterized by unorder residues [58], that could be able to interact with both G-protein-coupled receptors. In this context, intrinsically disordered peptides (IDPs) represent an innovative and interesting starting point to employ the conformational disorder in the design of potential drug agents in the after specific binding modifications. These peptides present a number of advantages with respect to structured peptides, such as a large interaction surface, involvement in high specificity/low affinity interactions, enhanced binding kinetics and conformational plasticity to bind several targets. Instead, they dynamically explore an ensemble of unfolded configurations, adopting more stable, ordered structures only after binding to their ligands. Two peptides with a high content of disorder promoting residues, PeptideE (H-YGECPCE-Allyl) and PeptideK (H-YGECPCK-Allyl) were designed using the N-terminal region of CXCL12 as template and the CPC sequence, the binding motif for CXCR4 receptor [51] that represents a well-known target for cancer therapies. The allyl group is introduced at C-*terminus* of both peptides in order to have a site for further new functionalizations. The peptides were analysed by solution CD and NMR techniques in order to study their conformational behaviour in solution. Fluorescence and DLS experiments were carried out in order to predict aggregates formation and to define their topology. Molecular simulation studies were conducted to verify if these peptides undergo structural changes upon binding or simply bind to pre-existing conformers that are populated within the IDP's structural ensemble. We have also perfomed biological tests to evaluate if PeptideE and PeptideK and their PA aggregates were able to interact with the CXCR4 receptor, to inhibit CXCL12-induced migration and to reduce cAMP monophosphate levels. The results of biological properties of all peptides were analalyzed with the purpose of comparing our compounds respect the most clinically advanced CXCR4 antagonist, i.e. plerixafor (previously known as AMD3100).

2.3. Results and discussion

2.3.1. Design and Synthesis

Two linear intrinsically disordered peptides have been designed choosing a few amino acids with disorder propensity: PeptideE (YGECPCE-Allyl) and PeptideK (YGECPCE-Allyl) (top of ***Figure 25***).

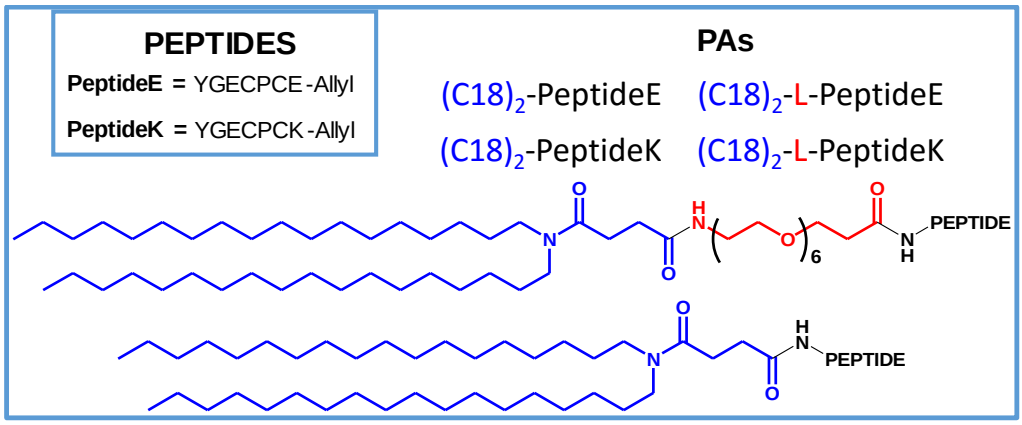


Figure 25*. Peptide sequences and their PAs derivatates.* ***Blue*** *for the hydrophobic moiety.* ***Red*** *fo the linker. In brackets the PEG chain.*

The charged amino acids, Lys and Glu, were introduced in order to evaluate the influence of charges in modulating the free peptide binding properties and on the aggregation tendency of their PAs derivates. As reported in ***Figure 25****, two* peptide amphiphiles (PA): $(C18)_2$-PeptideE and $(C18)_2$-PeptideK, were also obtained by conjugating two aliphatic chains (in blu) of eighteen carbon atoms at the N-*terminus* of both sequences. Two other PA derivatives, indicated in ***Figure 25*** as $(C18)_2$-L-PeptideE and $(C18)_2$-L-PeptideK, were obtained by inserting the AhOh exoethylene linker (red tract in ***Fugure 25***) between the peptide sequence and the hydrophobic double-tail. As it is well-known from the literature [62], the presence of PEG chains can increase the *in vivo* blood circulation of self-assembled PA aggregates. This linker, the same for both peptides, allows us to increase the hydrophilicity of PAs with out modifying their final charge. In this way, avoiding *intra-* and *inter-* electrostatic interactions, the linker leads to a distancing between the bioactive peptide and the hydrophobic shell of the aggregates. The relatively short length of the peptide sequence was chosen in order to avoid possible formation of a hydrophilic pocket, able to hide the amino acid sequence. PAs (***Figure 25***) were synthesized according to standard solid phase peptide synthesis protocols using 2-Cl-(Trt)-Cl preloaded resin as a polymeric support. (In **Material and Methods** section synthesis details are reported). The ethoxylic moiety and the two alkyl chains at the N-terminus of both peptide sequences were added as previously reported [63]. After RP-HPLC purification, the products were identified by ESI-MS spectrometry. The conformational features of both peptides were investigated in solution by CD, DLS, NMR, MD techniques.

2.3.2. Structural characterization

2.3.2.1. PeptideE and PeptideK

2.3.2.1.1. CD analysis

The secondary structure of **PeptideE** and **PeptideK** was studied in PBS by CD. As expected for very short peptide sequences, both peptides did not show any tendency to fold under this experimental condition [64]. Infact, CD spectra (***Figure 26***) show a typical shape of an unordered structure with a negative band between 197 and 201 nm.

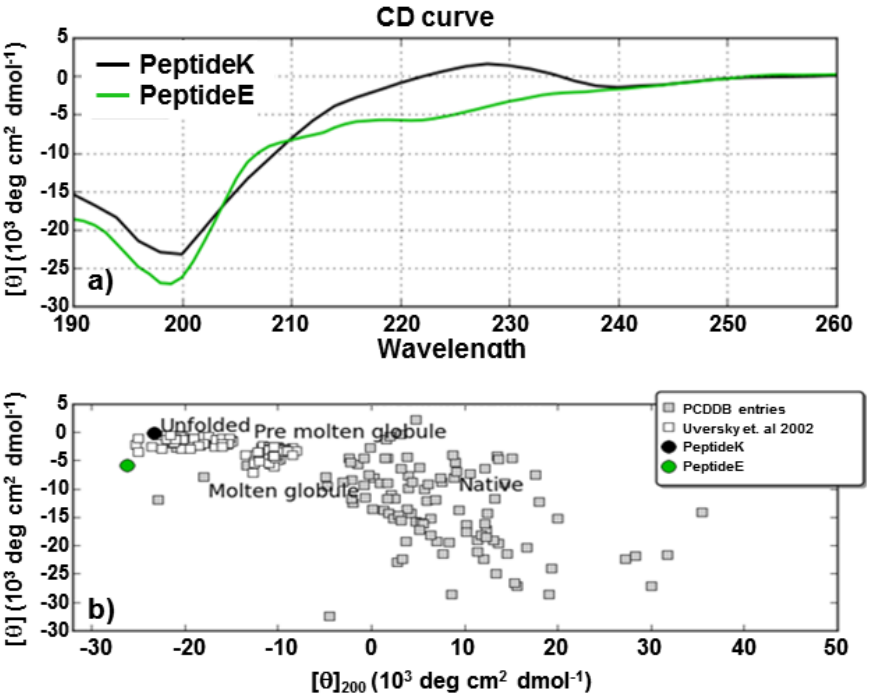


Figure 26 a) *CD spectra of PeptideK and PeptideE.* ***b)*** *CD analysis by CAPITO tool where both peptides are indicated as unfolded peptides.*

This behavior is also confirmed by CAPITO analysis [65] (http://capito.nmr.fli-leibniz.de), where spectra values (indicated as mean residue ellipticity (θ)) at λ = 200 nm are plotted *versus* λ = 222 nm. In fact, both the peptides are located in unfolded regions. (see ***Figure 26***) It is important to notice that the CD spectrum of PeptideK presents also a maximum centered at 225 nm that that is associated to the presence of fluctuating Polyproline II (PII) helices [66; 67, 68]. Conformational behaviours of PeptideE and PeptideK amphiphiles derivates ($(C18)_2$- and $(C18)_2$-L- derivates), were investigated by CD spectroscopy. Lyophilized peptides were dissolved in the suitable weakly basic buffer (TRIS or PBS, see paragraph **Material and Methods**) at concentrations above the CMC value (2×10^{-4} mol/L for all peptides) to foster the presence of aggregates in all solutions.

2.3.2.1.2. NMR Analysis

In aqueous solution, the poor spectral dispersion of the 1D [1H] spectra and the almost complete absence of signal in the 2D [1H, 1H] NOESY [69] experiments indicated that both free peptides were very flexible (***Figures A1*** and ***A2*** in the Appendix). However, complete proton resonance assignments were achieved by combined analysis of the 2D [1H, 1H] TOCSY [38] and ROESY [39] spectra (**Figures 27** and **28**; **Tables A1** and **A2**).

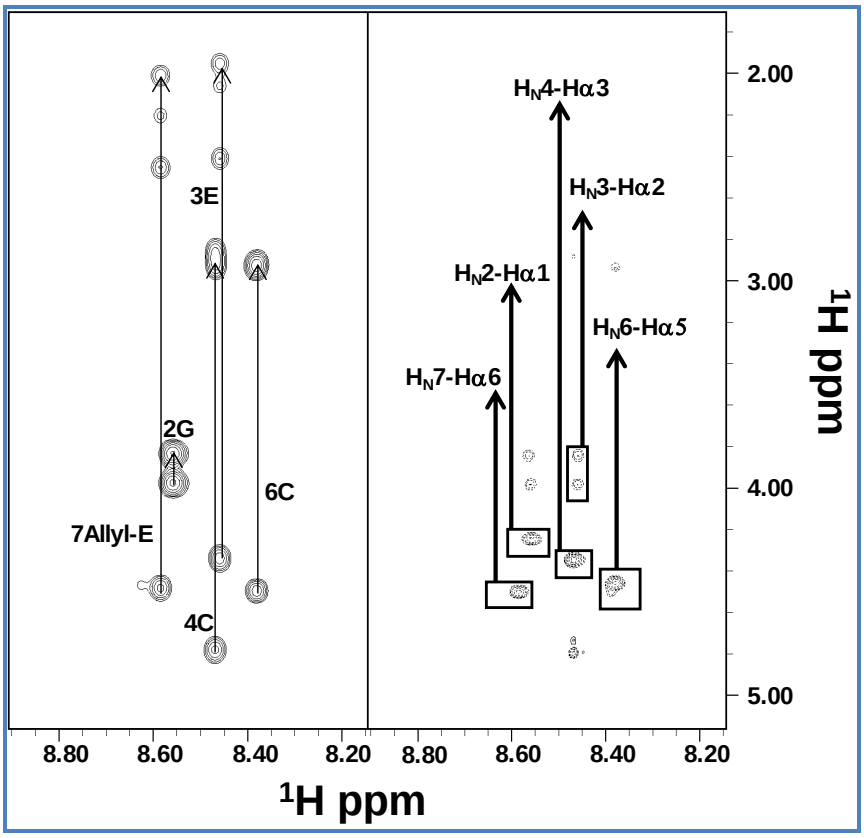


Figure 27. *Comparison of 2D [1H, 1H] TOCSY (**left**) and ROESY (**right**) spectra of **PeptideE** in H2O/D2O (90/10). The HN-aliphatic protons correlation regions are shown in each panel; spin system assignments are indicated in the left side. In the right panel, sequential ROE contacts are highlighted by rectangles and the corresponding assignments are indicated.*

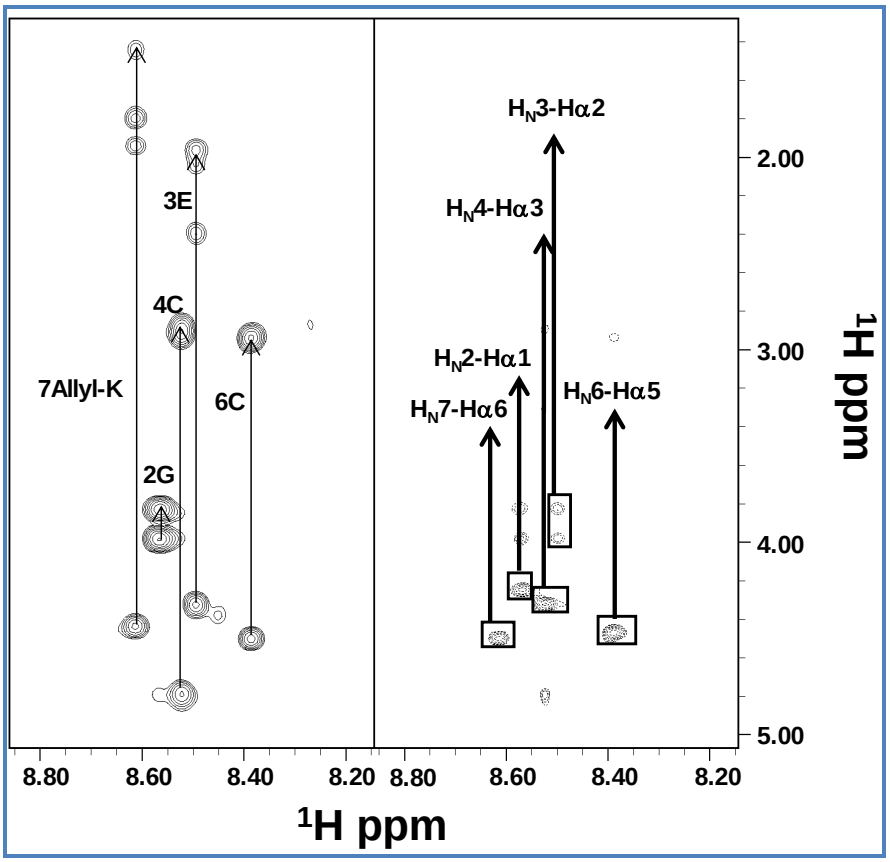


Figure 28. *Comparison of 2D [1H, 1H] TOCSY (**left**) and ROESY 200 (**right**) spectra of **PeptideK** in H2O/D2O (90/10). The two sides of the figure show spectral regions containing HN/aliphatic protons correlations; spin system assignments are indicated in the TOCSY left panel. In the right panel, sequential ROE contacts are shown.*

Moreover, we evaluated chemical shifts deviations of Hα protons from random coil values, which resulted in prevalence small and positive and characteristic of an extended disordered conformation [70]. (***Figures 29*** and ***30***, ***Tables A3*** and ***A4***)

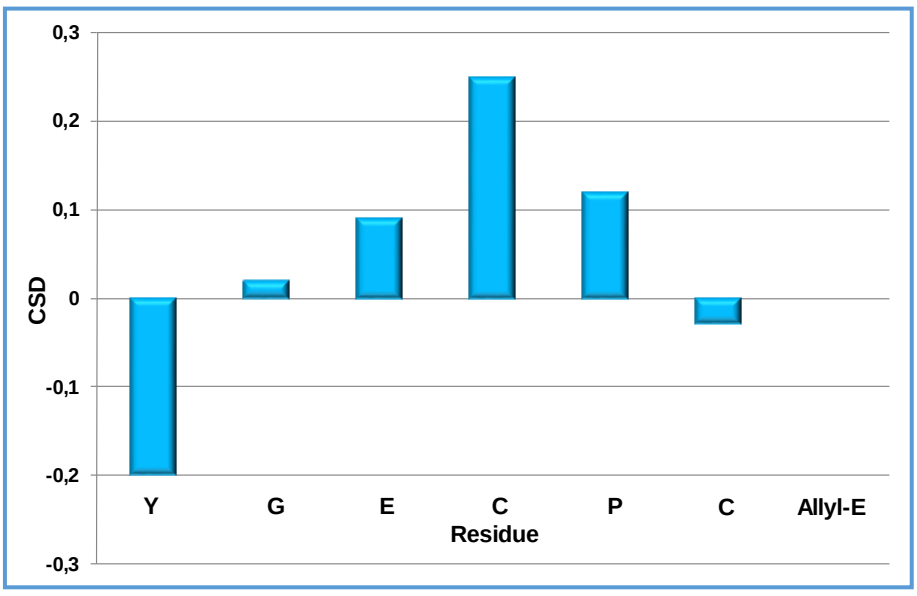


Figure 29. *Chemical shift deviations from random coil values (CSD) for PeptideE (D.S. Wishart et al., 1991). All residues are reported by one-letter code. The CSD value for Allyl-E is not shown.*

.

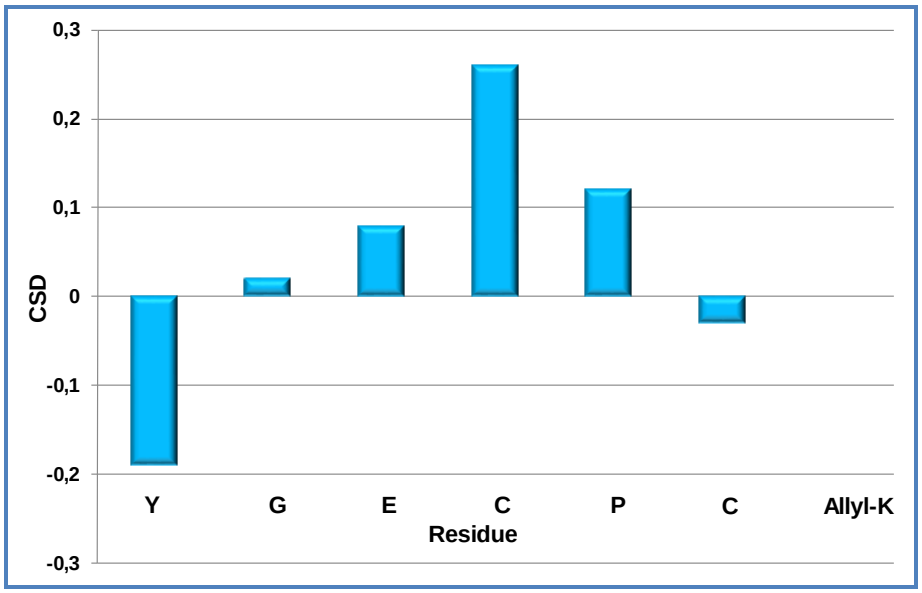


Figure 30. *Chemical shift deviations from random coil values (CSD) for PeptideK (D.S. Wishart et al., 1991). All residues are reported by one-letter code. The CSD value for Allyl-K is not shown.*

The disorder state of both peptides was confirmed by ROE patterns (***Figure 31***) which showed strong and sequential Hα*i*–HN*i+1* contacts typical of flexible peptides [71]. Due to the proximity of chemical shifts between Cys4 Hα proton and water in both PeptideE and PeptideK, we could not clearly identify the configuration (*i.e.*, *cis* or *trans*) of the Cys4–Pro5 peptide bond.

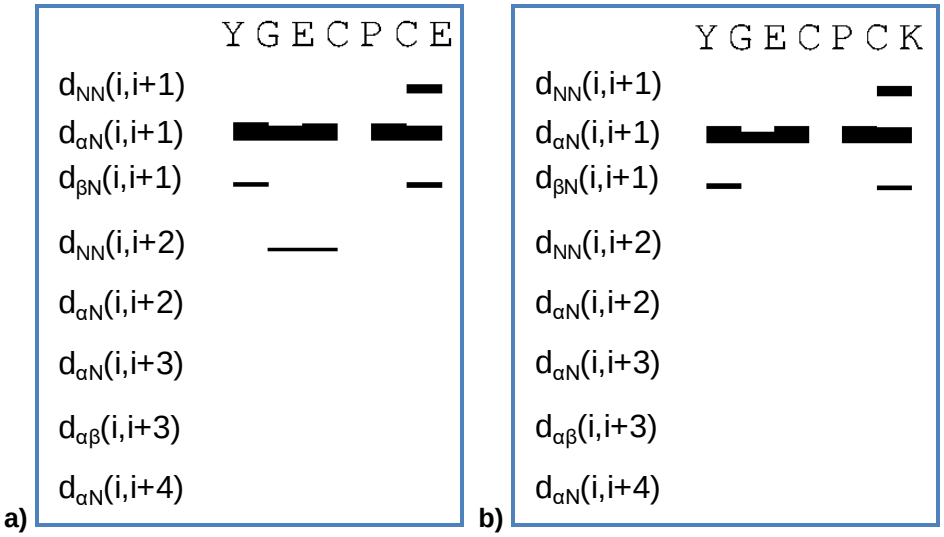


Figure 31. *NOE intensity pattern in H2O:* ***a)*** *PeptideE and* ***b****) PeptideK. Only principal short and medium range NOEs are reported. Different residues are specified with the one-letter amino acid code; "dξζ (b, c)" designates a NOE contact between the Hξ and Hζ protons in the b and c residues respectively.*

2.3.2.1.3. MD

Molecular dynamics studies were conducted on the two peptides, which were linearly modeled by using a precise method (**Material and Methods**), at neutral pH. On the basis of the distances between structures, like RMSD values, a set of clusters, reflecting the range of conformations accessible and their relative weights, was found by using a clustering algorithm implemented in GROMACS as reported also in recent papers [59; 66]**.** The plots of the Root Mean Square Deviation (RMSD) were computed by overlapping the various structures during simulations in respect to the initial conformation. In ***Figure 32a*** the RMSD plots of PeptideK and PeptideE reported. They evidenced high levels of fluctuation, thus suggesting that these peptides are flexible. This result was confirmed by Root Mean Square Fluctuation (RMSF) plots where the residues located at *N*- and *C*-termini have higher values of RMSF (***Figure 32b***). Instead, the gyration radii decreased during the simulations (***Figure 32c***) reaching a value of 0.5 nm. This suggests an increasing compactness of the two peptides, which result as if they were stabilized by a certain number of H-bonds of main chain–main chain (MM), main chain–side chain (MS) and side chain–side chain (SS) types (***Figure 32d***). This is the reason why the two peptides tend to become more compact and their radius of gyration to decrease (see ***Figure 32c***). Considering CD and NMR data in aqueous solution, the analysis of the secondary structure evolution clearly showed the high flexibility of both peptides during the simulations, which makes difficult the formation of stable regular secondary structure elements, *i.e.*, helix and β-strand.

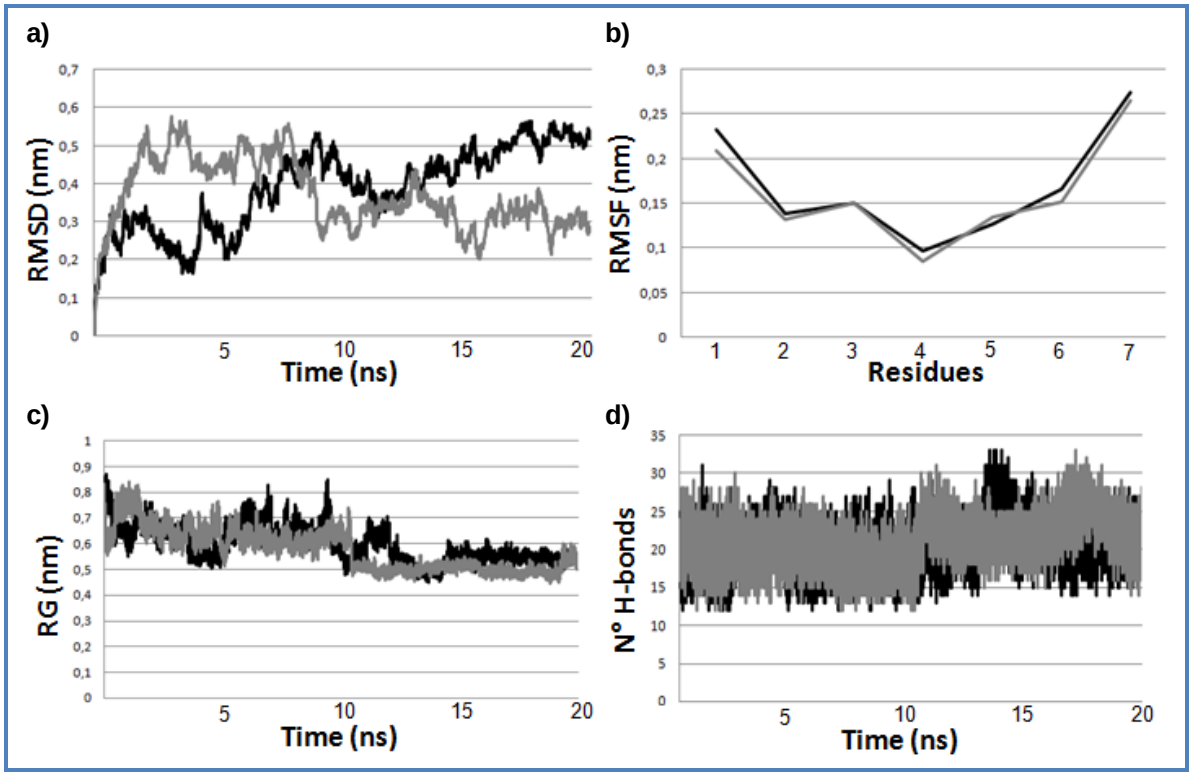


Figure 32*. Analysis of molecular dynamics simulations conducted on PeptideE (**in black**) and PeptideK (**in grey**) at physiological pH in terms of: **a)** root mean square deviation (RMSD) plot, **b)** root mean square fluctuation (RMSF), **c)** gyration radius plot, and **d)** H-bonds plot.*

A very flexible peptide structure is normally characterized by several conformers at equilibrium, hence they could be well described by a conformational ensemble.

Having taken this property into proper account, we have firstly performed a cluster analysis to determine the groups of structures that share similar conformational features according to their RMSD values. (see ***Figure 33***)

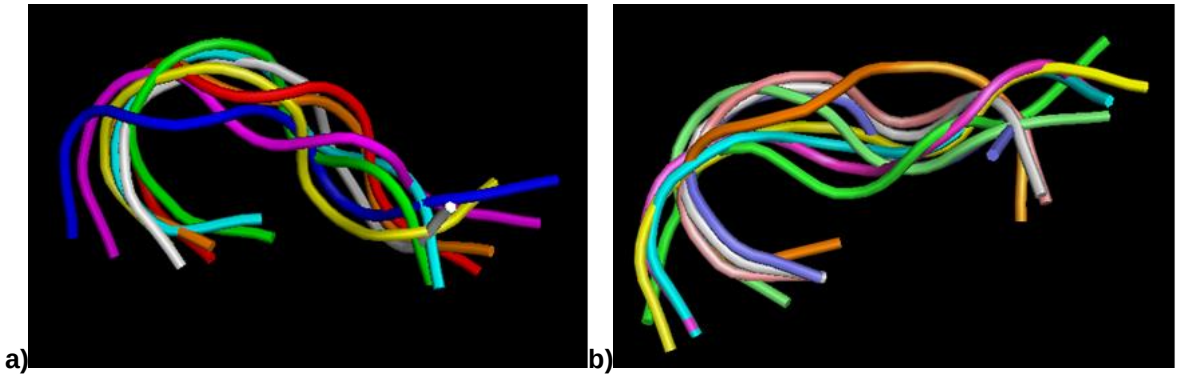


Figure 33. *Superimposition of* ***a)*** *8 clusters obtained for PeptideE and* ***b)*** *9 clusters for PeptideK during MD simulations.*

The most populated clusters are 8 for PeptideE and 9 for PeptideK. These clusters are mainly stabilized by MM H-bonds (***Figure 34***) that involved the atomic groups of the peptide backbone.

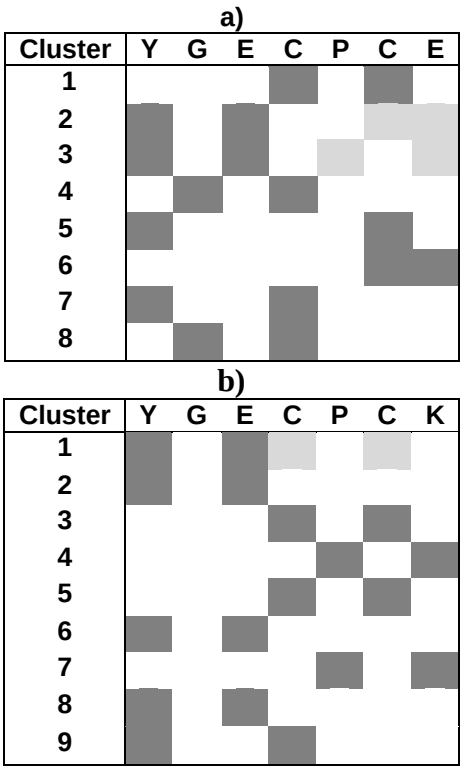


Figure 34. *The map MM H-bonds in the 8 and 9 clusters calculated for the* ***a)*** *PeptideE and* ***b)*** *PeptideK, respectevely, during MD simulations. We used the same colors the residues involved in the same MM H-bonds.*

This structural organization was observed by Pappu *et al.* [74] for charged peptides: it closely resembles the organization of collapsed and slightly soluble globules.
Since CD spectra of PeptideK pointed out the presence of PII, we verified its possible occurrence in the most populated clusters by using the angular ranges following reference parameters for the angles: $-110° \leq \varphi \leq -40°$ and $130° \leq \psi \leq 180°$ in the

Ramachandran map [75]. No residues in PII were present in eight clusters of PeptideE whereas in 5 out of 9 clusters of PeptideK we identified residues in PII (**Figure 35**), according to CD analysis.

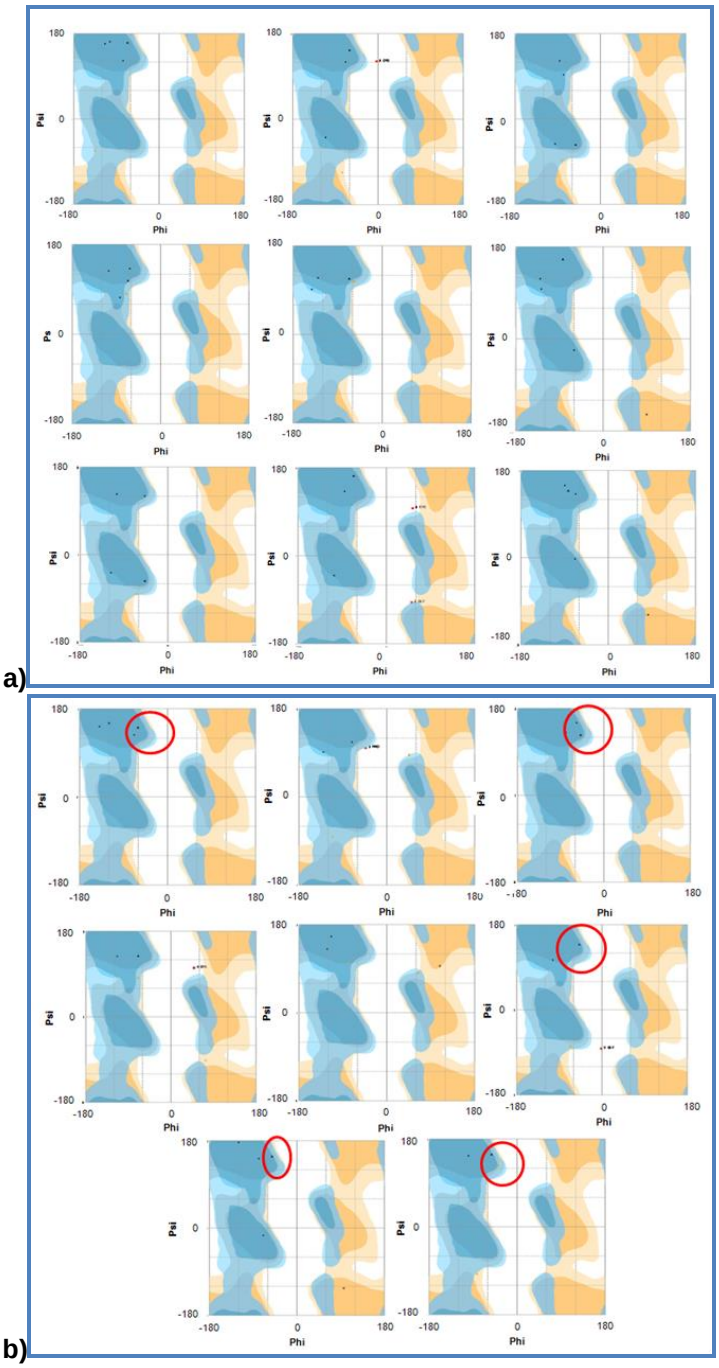


Figure 35*. Ramachandran Plot for nine and eight clusters obtained for* ***a)*** *PeptideE and* ***b)*** *PeptideK, respectively.* ***Red circle*** *for the the residues in polyproline II region.*

It is well known that the peptide dynamics depend on the surrounding solvent, which mediates interactions among residues. The total number of water molecules in the system was analysed as well as their possible role [76]: the simulation box for PeptideE and PeptideK contained 2956 and 2611 water molecules, respectively. The total average number of water molecules that formed H-bonds resulted slightly higher for PeptideE if compared to PeptideK: 33 and 30 in PeptideE and PeptideK,

respectively. This trend could be due to the different amino acid sequence of the two peptides; in fact, PeptideK has got two oppositely charged residues in 3 and 7 positions, hence it may assume a more compact structure. In both peptides most of H-bonds (indicatively 60%) resulted to involve the peptide backbone, being of main chain—water oxygen type (MH) or of the water oxygen and main chain type (HM). A detailed analysis of the residues involved in H-bonds with water molecules revealed that the residues in positions 1, 3 and 7 (Y, E and E in PeptideE and Y, E and K in PeptideK) participate to a higher number of H-bonds with water molecules if compared with the residues located in the other four positions (***Table 6***).

Table 6. *The average number of H-bonds that each residue in PeptideE and PeptideK formed with water molecules during MD simulations.*

Residue	PeptideE	PeptideK
Y	7	8
G	3	2
E	9	6
C	2	1
P	1	3
C	2	4
E or K	9	6

Furthermore, the average number of H-bonds developed by E and K with water molecules was lower than the average number of the H-bonds developed by E and E with water molecules in PeptideE. Hence, these results confirmed that PeptideE tended to assume more extended conformations being that it contains two negatively charged residues that could allow large interactions with water molecules.
Taking into account all data the structure of both peptides is very flexible, and can be dynamically stabilized by a network of MM H-bonds. Hence, their structures can be adequately represented as conformational ensembles characterized by fluctuating irregular secondary structures. (***Figure 34***) In fact, the peptides do not assume stable and fixed conformations: they go continuously from one cluster to another. The calculation a 3D model based on NMR parameters was hampered by the high flexibility of peptides. For this reason, the MD was very useful to understand how these peptides can move as well as to understand that the best conformers needed to be put together to have a clear representation. In particular, the two peptides do not explore the full space, but rather, the rapid interchange among the flexible conformers induce quick changes between clusters. This is coherent with the data reported in Figure 4B where it is visible that the more flexible residues are located in the terminal ends.

2.3.2.2. PAs derivates

2.3.2.2.1. Aggregates preparation and Fluorescence

Supramolecular aggregates of pure PA derivatives were obtainedby dissolving the lipophilic derivatives in 10 mM phosphate buffer (pH 7.4) for PeptideE and in 10 mM TRIS buffer (pH 8.0) for PeptideK. The choice of buffers to be used for solubilizing monomers was imposed by their different propensities to aggregate. All the aggregates properties were investigated by fluorescence spectroscopy and DLS. Critical micellar concentration (CMC) values were estimated by fluorescence spectroscopy by using ANS as a fluorescent probe. The fluorescence intensity of ANS depends on thesurrounding environment. Infact this fluorophore emits only in a hydrophobic environment, such as the hydrophobic core of amicellar aggregate, whereas it does not emit in water solution. CMC values can be estimated break-point in the graphics reported in ***Figure 36***.

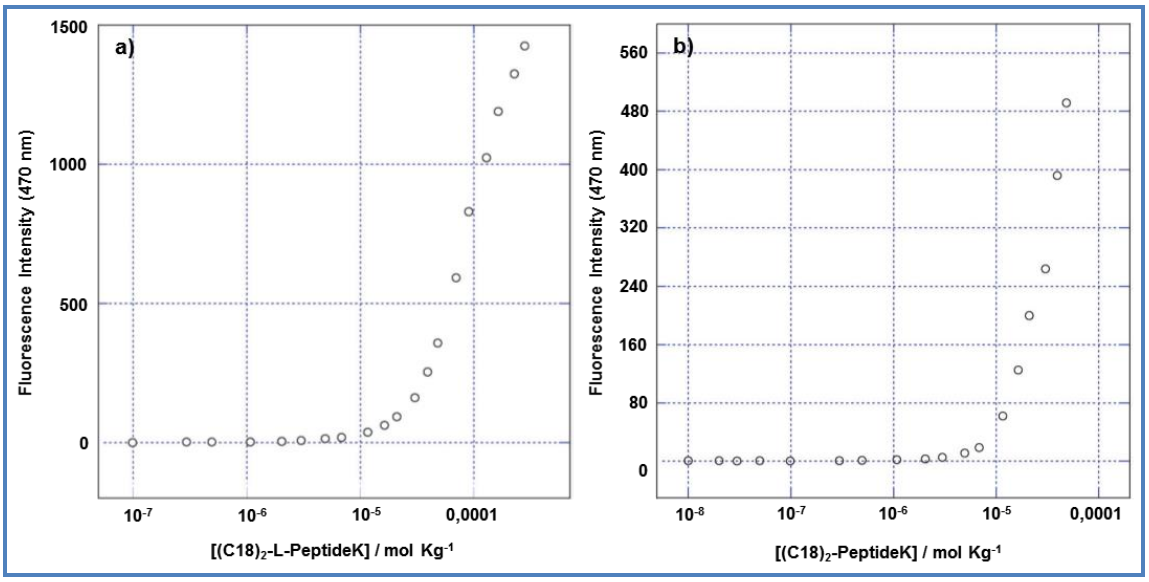


Figure 36*. Fluorescence intensity of the ANS fluorophore at 470 nm as a function of* ***a)*** *$(C18)_2$-L-PeptideK and* ***b)*** *$(C18)_2$-PeptideK concentration; data are multiplied by a scale factor for a better comparison. CMC values are obtained from graphical break points.*

These data have been obtained by plotting the fluorescence intensity of ANS in the emission maximum at 470 nm, as a function of the PA derivative concentration. The CMC values (7.0 x 10^{-6} mol kg^{-1} and 1.9 x 10^{-5} mol kg^{-1}, respectively for $(C18)_2$-PeptideK and $(C18)_2$-L-PeptideK) indicate that the introduction of the ethoxylic linker slightly affects the formation of aggregates capable of capturing the ANS, but they also confirm the high stability of both the resulting aggregates. (***Table 5***) The same trend was found for PeptideE. Nonetheless, the CMC values of both the amphiphilic derivatives of PeptideE are higher than the corresponding ones for PeptideK. (***Table 5***)

***Tabella 5**. Structural parameters for the aggregates obtained by fluorescence measurements (critical micellar concentration value, CMC) and by dynamic light scattering measurements (diffusion coefficients, D, hydrodynamic radii, RH, and polydispersity indices).*

Peptide	CMC (mol kg^{-1})	R_H (nm)	D x 10-12 (m^2s^{-1})	PDI
$(C18)_2$-PeptideK	7.0×10^{-6}	80 ± 27	3.1 ± 1.0	0.270
$(C18)_2$-L-PeptideK	1.9×10^{-5}	82 ± 34	3.1 ± 1.5	0.182
$(C18)_2$-PeptideE	3.9×10^{-5}	112 ± 44	1.9 ± 0.7	0.264
$(C18)_2$-L-PeptideE	3.2×10^{-5}	110 ± 39	2.4 ± 0.5	0.226

This behavior indicates that PeptideE owns a lower propensity to aggregate if compared to PeptideK; this may depend on the different interaction of the charges with the buffer salts.

2.3.2.2.2. DLS

DLS data (hydrodynamic radii, RH; the diffusion coefficients, D; and the polydispersity indices, PDI) are reported in ***Table 5***. The measurements for the self-assembled PAs were performed at y = 1731 at a concentration of 2 x $10^{-}4$ mol $kg^{-}1$ (above the CMC values of all amphiphilic peptides) in the previously reported buffer solution. The Stokes–Einstein equation **(1)** is used to evaluate the hydrodynamic radius (RH) at infinite dilution where $\boldsymbol{D}_0$ is the translational diffusion coefficient at infinite dilution, $\boldsymbol{K}_B$ is the Boltzmann constant, ***T*** is the absolute temperature, and $\boldsymbol{\eta}$ is the solvent viscosity. Due to the high solution dilution (C = 1 x 10^{-4} M) of the systems under study, we have approximately $\boldsymbol{D} \sim \boldsymbol{D}_0$, and equation (1) can be reasonably used to estimate the hydrodynamic radius of the aggregates.

$$R_H = \frac{K_B T}{6\pi\eta D_0} \quad \textbf{(1)}$$

Taking into account all DLS results, it was possible to oberve that all aggregate solutions show a monomodal distribution, which indicates the presence of just one population of aggregates with an average diameter ranging between 80 and 110 nm. An increase of the radius of 20% was observed for $(C18)_2$-PeptideE (100 nm) with respect to $(C18)_2$-PeptideK (80 nm) (see ***Figure 37***). Whereas only a slight difference in radii can be detected for the derivatives containing the spacer. (***Table 5***)

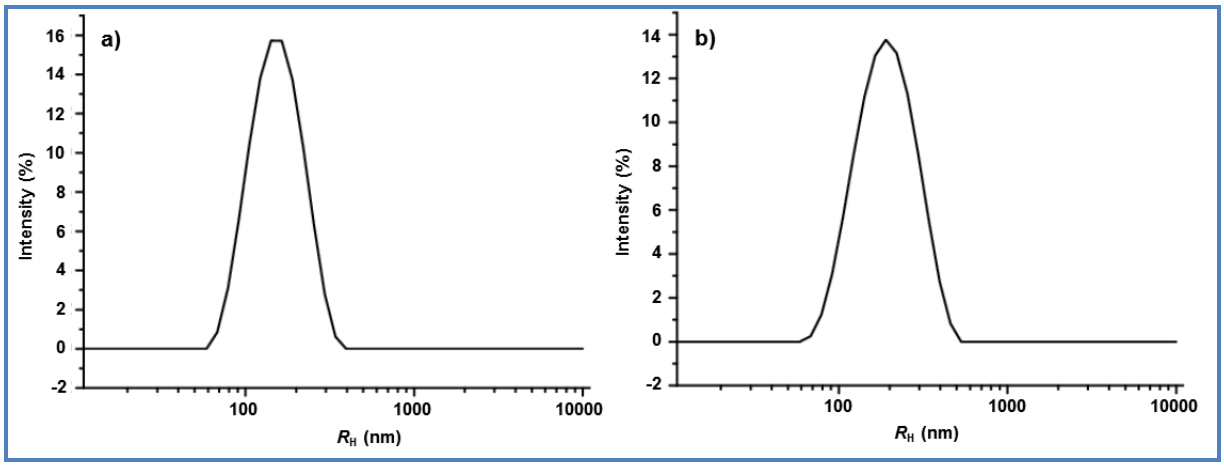
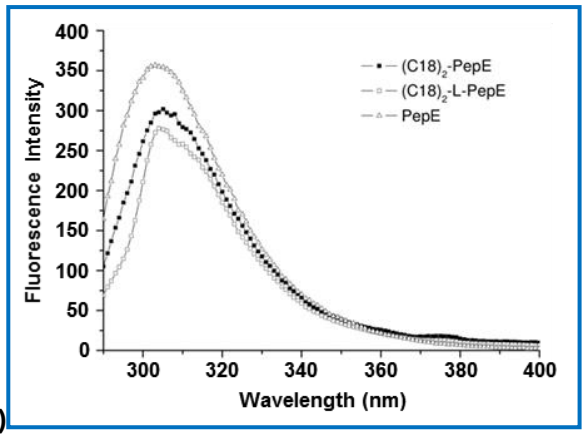


Figure 37. *DLS profiles of $(C18)_2$-PeptideK* ***a)*** *and $(C18)_2$-PeptideE* ***b)*** *peptide amphiphiles at a concentration of 2 x 10^{-4} M.*

Fluorescence measurements were carried out by exciting buffer solutions at 275 nm to evaluate the eventual shifts of tyrosine fluorophore emission as a consequence of aromatic quenching on the surface of the aggregates. In the ***Figure 38a*** emission spectra for all PeptideE derivatives are reported.

a) 

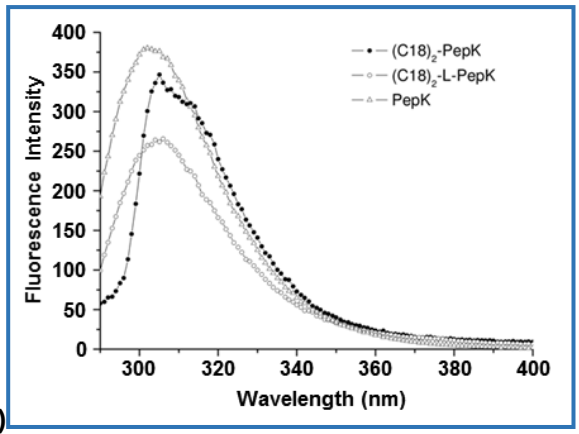
b) 

***Figure 38**. Fluorescence emission spectra recorded by exciting buffer solutions at 275 nm for **a)** PeptideE derivatives and **b)** PeptideK derivatives.*

The behaviour of tyrosine does not show significant differences in intensity or in the wavelength shift. These results seem to exclude interaction among peptide side chains on the surfaces of aggregates. Similar results have been obtained for PeptideK derivatives (***Figure 38b***).

2.3.2.2.3. CD analysis

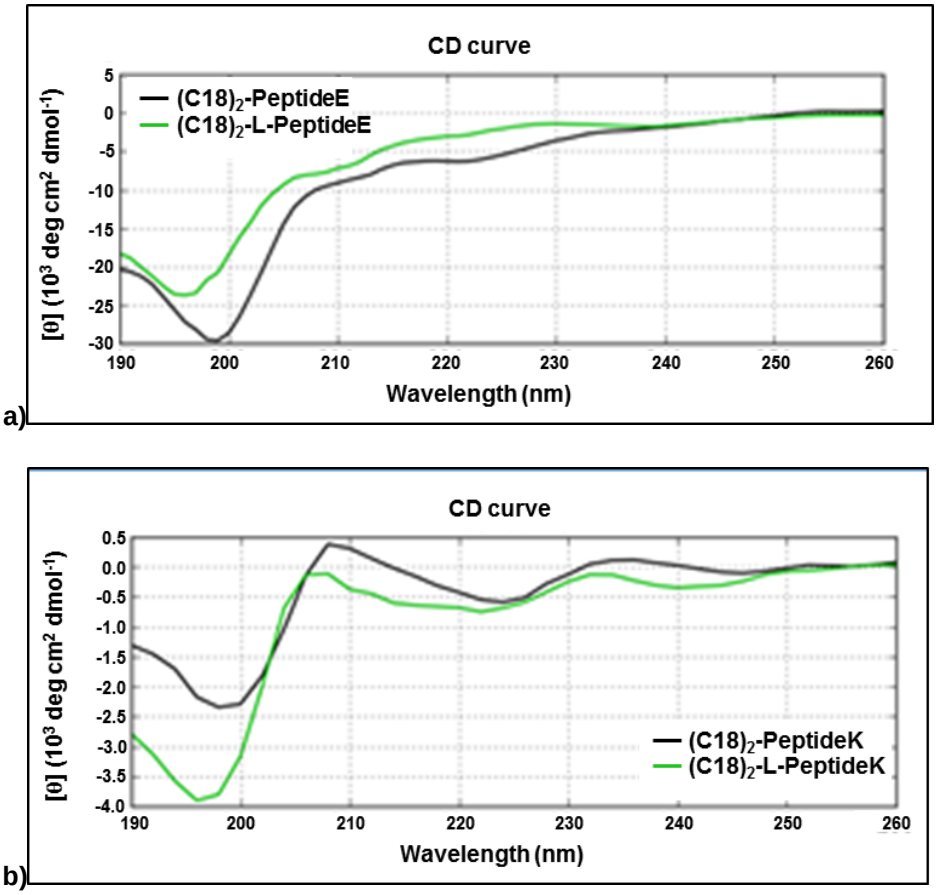


***Figure 39**. CD spectra of: **a)** (C18)2-PeptideE and (C18)2-L-PeptideE; and **b)** (C18)2-PeptideK and (C18)2-L-PeptideK.*

In ***Figure 39*** the CD spectra of all PA derivates are depicted. The CD spectrum of $(C18)_2$-PeptideE does not show any tendence to fold as indicated by CD spectra, presenting a negative minimum at around 198 nm, typical for unordered structures (***Figure 39a***). Moreover, the presence of the ethoxylic spacer does not influence significantly the dichroic bent of the PA. This trend was also confirmed by CAPITO analysis (data not shown) [65]. Furthermore, two weak maxima at 208 and 232 nm are also detectable in the CD spectra of PeptideK derivatives (see ***Figure 39b***), in addition to the minimum at 196 nm. These spectra suggest that the replacement of the glutamic acid with the lysine at the C-terminus induces a rearrangement of the secondary structure with a partial folding of the IDP.

2.3.2.2.4. NMR Analysis

The high aggregation levels of PeptideE and PeptideK derivatives prevented any attempts to confirm the CD data with more precise conformational studies by NMR spectroscopy. Indeed, this aggregation produced sample precipitationand/or complete disappearance of signals in 1D [1H].

2.3.3. Functional Characterization on CXCR4 Receptor

All peptides have been tested in order to otain different information about our peptide activity on the CXCR4 receptor. In order to assess the inhibitory capability of PeptideK and PeptideE on CXCR4 function, we have evaluated peptides binding to the receptor through an indirect binding assay; then, in addition we have assessed the inhibition of CXCL12-induced migration and, considering the correlation between the CXCR4 and adenylate cyclase activities (***Figure 40***), the percentage of cAMP. The binding of the four peptides to CXCR4 was evaluated in CCRF-CEM using a CXCR4 12G5 monoclonal antibody as previously described, [56] and two concentrations (1 mM and 10 mM) which are matching the ones previously used in CXCR4 binding assays. All peptides did not displace 12G5 antibody if compared with AMD3100, which strongly reduced antibody binding to CXCR4. This result could be justified by the presence of different binding sites on AMD3100/12G5 antibody and the four highly flexible disordered peptides. The migration assay revealed that PeptideE and PeptideK reduce the migration of CCRF-CEM cells towards CXCL12 even if they showed less activity than AMD3100 (data not shown). Similar results were obtained for PAs derivates, indicating the absence of influence of the hydrophobic and/or linker on the migration. As a consequence of the activity of CXCR4 on the G-protein (mediated by CXCL12), the adenylate cyclase is inhibited. For this reason, we also evaluated the cAMP levels after treatment with the all peptides. The results indicate that in the presence of CXCL12 (100 ng mL^{-1}) and Forskolin (1 mM):

- **i.** All synthetic peptides show a dose dependent increase of the cAMP (average value of 66% anf 84% at 1 mM and 10 mM), respectively.
- **ii.** At 10 mM, $(C18)_2$-L-PeptideK increases cAMP levels up to 93% and $(C18)_2$-PeptideE increases the cAMP level to 95%;
- **iii.** AMD3100 exhibits less efficient activities on the adenylate cyclase when compared to the four peptides (***Figure 36***).

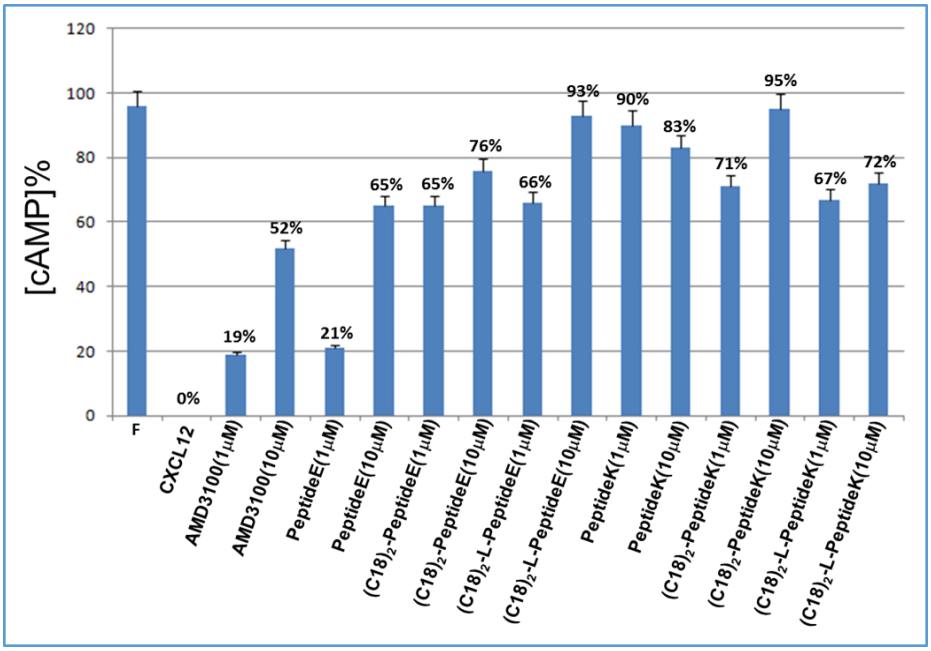


***Figure 40**. Comparison of cAMP modulation by PeptideE, PeptideK, $(C18)_2$-PeptideE, $(C18)_2$-L-PeptideE, $(C18)_2$-PeptideK, $(C18)_2$-L-PeptideK and AMD3100 at 1 μM and 10 μM.*

2.4. Conclusions

As CXCR4 is overexpressed in several human cancers, the blockade of CXCR4–CXCL12 interaction represents from many years an interesting drug-target [51]. Nowadays, the most clinically advanced compound targeting this route, plerixafor (previously known as AMD3100), shows cardiotoxicity, as reported in its clinical trial against HIV [77]. From this, the need to design e synthesize new antagonists able to block, without side effects, the binding between CXCL12 and CXCR4. We started to study the structural preferences in solution of two synthetic linear peptides, PeptideE and PeptideK, characterized by few amino acids with disorder propensity, and by an allylic ester group at the *C*-terminal end. NMR and CD spectra evidenced the absence of regular secondary structure elements thus confirming the natively disordered nature of these peptides. MD simulations confirmed their flexibility but also revealed that they are stabilized by a network of transient and dynamic *intra*-molecular H-bonds of MM type and by interactions with water molecules. In conclusion, NMR and computational studies pointed out to us the flexibility of PeptideE and PeptideK, their inability to adopt a single ordered conformation and, hence, the need to represent them by conformational ensembles characterized by the presence of transient and dynamic MM H-bonds. Afterwards, we have explored the possibility that intrinsically disordered peptides could be connected to alkyl chains to generate new molecular buildings for drug delivery vehicles in cells over-expressing the CXCR4 receptor. The peptides are used as polar heads of these buildings whereas alkyl chains as a hydrophobic moiety. In detail, we investigated the opportunity to modulate conformational changes by introducing C18 chains and/or an ethoxylic linker. DLS and fluorescence data revealed that the insertion of aliphatic moieties into peptides favors the self-assembling in supramolecular aggregates of the resulting amphiphilic molecules. The average diameter ranges between 80 and 110 nm, independently from their charge and from the presence of the PEG spacer. Once confined on the external surface of the aggregates, peptides could undergo a disorder-to-order transition [57]. The low solubility at high concentrations (0.50–1.0 mM) and the strong aggregation of all PAs hampered their structural characterization by NMR techniques, even after changing the experimental conditions such as implementing buffers at different pH and temperatures. CMC values agree with those previously reported for amphiphilic peptides with the same hydrophobic moiety and they pointed out the high stability of the aggregates [78; 79]. The lower CMC values of PeptideK PAs derivates indicate a higher propensity to aggregate if compared with the corresponding PeptideE derivatives. Furthermore, the insertion of the hydrophobic portion in PeptideE seem not to influence its tendency towards the disordered/unfolded state from the conformational point of view. Whereas both the PeptideK derivatives show a partial folding, with respect to the free peptide sequence. To the best of our knowledge, these compounds are the first CPC motif containing amphiphilic peptides which are able to aggregate. Comparing all peptides with AMD3100, the studies of functional characterization on CXCR4 Receptor, show *no significant* binding affinity of all peptides for CXCR4. Concerning migration assays, the best results about the migration were obtained for PeptideE sequences, indicating the importance of the choice of charge in the desing of bioactive peptide. As indicated by the highest cAMP percentages of $(C18)_2$-L-PeptideE-(93%) and $(C18)_2$-PeptideK (95%), obtained at 10 mM (***Figure 36***), the presence of the linker could influence the activity of PAs. Although these results should be considered as preliminary, requiring further analysis, we can assume that all data obtained for the amphiphilic peptides designed and synthesized were encouraging and indicate the

validity of multidisciplinary approach described in this work. Further chemical modifications using the reactivity of the ester group could be introduced to obtain new molecules able to form supramolecular aggregtes peptide-based, capable of being encapsulated by living cells and, therefore, convey the bioactive peptide and / or probes. All this can find applications in the field of medical therapies personalized, investigations MRI (magnetic resonance imaging) and also for biotechnological industry.

2.5. Material and Methods

2.5.1. Reagents

Protected fluorenylmethoxycarbonyl (Fmoc)N $^{\alpha}$-Fmoc-amino acid derivatives, 2-chlorotrityl chloride resin pre-loaded with Fmoc-Glu-OAllyl and Fmoc-Lys-OAllyl and coupling reagents were purchased from Novabiochem, Iris Biotech GMBH or from Inbios (Napoli, Italy). All other chemicals were commercially available by Sigma-Aldrichand Acros. Fmoc-21-amino-4,7,10,13,16,19-hexaoxaheneicosanoic acid (Fmoc-Ahoh-OH) was purchased from Neosystem (Strasbourg, France). N,N-Dioctadecylsuccinamic acid was synthesized according to the published methods [80]. The purity of these products was determined by analytic reverse-phase HPLC using a VWR Hitachi instrument equipped with an L-2450 auto sampler, two L-2130 pumps, a Satisfaction RP18-AE column (5 µm, 250 × 4.6 mm) and a L-2450 diode array detector, at a flow rate of 0.80 mL/min. The free peptides were purified by preparative reverse-phase HPLC using a VWR LaPrep system consisted of a P202 injector, two P110 pumps, a Satisfaction RP18-AB C18 column (5 µm, 250 × 20 mm) and a P314 UV detector, at a flow rate of 10.0 mL/min. The following eluents were used in a gradient mode: (A) 0.1% trifluoroacetic acid (TFA) in H_2O/CH_3CN (95/5) and (B) 0.1% TFA in CH_3CN/H_2O (95/5). Water was of Milli-Q quality and was obtained after filtration of distilled water through a Milli-Q® cartridge system. CH_3CN and TFA were of HPLC use quality. Degassing of solvents was performed using argon bubbling. Gradient used for the analytic and preparative RP-HPLC was respectively from 0% B to 100% B in 15 min and from 0% to 70% B in 70 min; UV detection at 214 nm. CH_3CN was evaporated and the aqueous solution was freeze-dried to give purified PeptideK and PeptideE as white solids (68% and 54% yield, respectively). Peptide identity was confirmed by mass spectrometry analyses performed on an Ultraflex III TOF/TOF system (Bruker Daltonics, Bremen, Germany), equipped with 200 Hz smartbeam laser (355 nm) and operating in reflectron positive ion mode. Mass spectra were acquired over the m/z range 450–5000 by accumulating data from 1000 laser shots for each spectrum. The instrumental conditions employed to analyze molecular species were the following: ion source 1: 25.08 kV; ion source 2: 21.98 kV, lens: 11.03 kV, pulsed ion extraction: 30 ns, reflector: 26.39 kV, reflector 2: 13.79 kV. Matrix suppression was activated by deflection mode: suppression up to 450 Da. Mass calibration was performed for each samples in range of ~700 Da–3200 Da with a peptide calibration mixture (8206195, Peptide Calibration Standard, Bruker Daltonics). The instrument was controlled using Bruker's flexControl 3.4 software and mass spectra were analyzed in Bruker's FlexAnalysis 3.4 software. Concerning PAs derivates, all molecules were purified on on a LC8 Shimadzu HPLC system (Shimadzu Corporation, Kyoto, Japan) equipped with a UV lambda-Max Model 481 detector using a Phenomenex (Torrance, CA) C4 (300 Å, 250 x 21.20 mm, 5 µ) column. Elution solvents are H_2O/0.1% TFA (A) and CH_3CN/0.1% TFA (B), from 20% to 95% over 20 minutes at 20.0 $mLmin^{-1}$ flow rate. The purity and identity of the products were assessed by analytical LC–MS analyses by using Finnigan Surveyor MSQ single quadrupole electrospray ionization (Finnigan/Thermo Electron Corporation San Jose, CA), column C4-Phenomenex eluted with an H_2O/0.1% TFA (A) and CH_3CN/0.1% TFA (B) from 20% to 95% over 20 minutes at 0.80 mL min^{-1} flow rate.

2.5.2. Synthesis

All peptides were synthesized following a solid phase peptide standard protocol using Fmoc chemistry on 2-chlorotrityl chloride resin pre-loaded (0.24 scale mmol) with the side chain of a Fmoc-Lys-OAllyl and a Fmoc-Glu-OAllyl, respectively. The sequence H-YGECPC was built on both resins. In particular, the resins were allowed to swell in dichloromethane (DCM) and Fmoc-deprotection steps were performed by treating the resin for 30 min in a dimethylformamide (DMF)/Piperidine (Pip) (80/20, v/v) solution. For coupling steps the resin was suspended in DMF containing three equivalents of the Fmoc-protected amino acid, O-benzotriazole-tetramethyl-uronium-hexafluoro-phosphate (HBTU) (3eq), and diisopropylethylamine (DIEA) (6 eq.), and the reaction mixture was shaken for 2 h at room temperature. Washing steps after coupling and deprotection steps were performed by treating resin in DMF (3 × 1 min), MeOH (1 × 1 min), and DCM (3 × 1 min), successively. The Fmoc ethoxylic spacer and lipophilic N,N-dioctadecylsuccinic acid were coupled, by using a procedure already reported in literature [63]. Peptides were cleaved by soaking the resin in TFA/H_2O/triisopropylsilane (TIS) (3.00 mL, 95/2.5/2.5, v/v/v) for 3 h. The peptide was cleaved by immersing the resin in TFA/H2O/triisopropylsilane (TIS) (30.0 mL, 95/2.5/2.5, *v*/*v*/*v*) for 3 h. The resulting white solids were dissolved in a water and acetonitrile 4 : 1 v : v (10.0 mL) mixture and freeze-dried to obtain a white powder that was analyzed for purity and purified by preparative RP-HPLC. Mass values were obtained by LC-ESI-MS (see ***Table 7***).

Table 7. *Experimental masses of Pas derivates.*

Peptide	Experimental Mass
$(C18)_2$-PeptideE	1443 a.m.u.
$(C18)_2$-PeptideK	1442 a.m.u.
$(C18)_2$-L-PeptideE	1777 a.m.u.
$(C18)_2$-L-PeptideK	1776 a.m.u.

2.5.3. Preparation of aggregate solutions

Lyophilized PAs were dissolved in aqueous solutions. In detail, $(C18)_2$-PeptideK and $(C18)_2$-L-PeptideK were solubilzed in 10 mM TRIS buffer at pH 8.0, whereas $(C18)_2$-PeptideE and $(C18)_2$-L-PeptideE in 10 mM phosphate buffer (PBS) at pH 7.4. The calibration of the pH-meter was carried out by using three standards (pH 4.00, pH 7.00 and pH 10.00). In most cases the samples to be measured were prepared from stock solutions. The concentrations of all solutions were determined by UV-Vis absorbance measurements carried out on a Thermo Fisher Scientific Inc (Wilmington, Delaware USA) Nanodrop 2000c spectrophotometer equipped with a 1.0 cm quartz cuvette (Hellma) using a molar absorptivity (ε) of 1390 M^{-1} cm^{-1} at λ = 275 nm, which corresponds to the wavelength of the tyrosine residue.

2.5.4. Fluorescence measurements

CMC values of PAs derivates were obtained by fluorescence measurements. Fluorescence spectra were recorded at room temperature on a Jasco Model FP-750 spectrofluorophotometer (Jasco Int. Co. Ltd, Tokio, Japan) in a 1.0 cm path length quartz cell. Equal excitation and emission bandwidths were used throughout the experiments, with a recording speed of 125 nm min^{-1} and automatic selection of the time constant. The 8-anilino-1-naphthalene sulfonic acid ammonium salt (ANS) was

used as fluorescent probe [80]. Small aliquots of peptide aqueous solution were added to a fixed volume (1.00 mL) of 2.0 x 10^{-5} M ANS fluorophore directly in the quartz cell. CMC values were determined by linear least-squares fitting of the fluorescence emission at 480 nm, upon excitation at 350 nm versus the amphiphile concentration. Fluorescence emission spectra of the tyrosine residue in the anphiphilic peptides were recorded at 5 x 10^{-5} M concentration and 25 °C exciting the peptide solution at 275 nm.

2.5.5. DLS

Dynamic light scattering (DLS) measurements were carried out using a Zetasizer Nano ZS (Malvern Instruments, Westborough, MA) that employs a 173° backscatter detector. Other instrumental settings are measurement position (mm): 4.65; attenuator: 8; temperature 25 1C; cell: disposable sizing cuvette. DLS samples were prepared at a final concentration of 2.0 10^{-4} M and centrifuged at room temperature at 13 000 rpm for 5 min.

2.5.6. CD

Far-UV CD spectra were recorded from 190 to 260 nm on a Jasco J-810 spectropolarimeter equipped with a NesLab RTE111 thermal controller unit using a 1 mm quartz cell at 25 °C. Circular dichroism measurements were carried out on lyophilized peptides dissolved in 10 mM phosphate buffer at pH 7.4 M at a peptide concentration of 5·10^{-6} M. CD spectra of PAs derivatives of PeptideE and PeptideK were acquired at 2.0 x 10^{-4} M concentration in 2.0 mM phosphate buffer and 10 mM TRIS, respectively. Other experimental settings were: scan speed, 10 nm min^{-1}; sensitivity, 50 mdeg; time constant, 16 s; bandwidth, 1 nm. Each spectrum was obtained averaging three scans, and by subtracting contributions from other species in solution and converting the signal to mean residue ellipticity in units of deg cm^2 $dmol^{-1}$ res^{-1}.

2.5.7. NMR Analysis

NMR spectra were recorded at 25 °C on a Varian Unity Inova 600 MHz spectrometer provided with a cold probe. The process of proton resonance assignments was carried out with a canonical protocol [71] based on analysis of the following two dimensional [^{1}H, ^{1}H] spectra: TOCSY (70 ms mixing time) [38], DQFCOSY [40], NOESY [69] (300 ms mixing time) and ROESY (Rotating frame Overhauser Enhancement Spectroscopy)) [39] (200 ms mixing time). Chemical shifts were referenced with respect to the TSP, 99% d, Armar Scientific, Switzerland) signal at 0.0 ppm. 1D spectra were acquired with a relaxation delay of 1s and 32-128 scans. 2D experiments were generally acquired with 32-64 scans, 128-256 FIDs in t_1, 1024 or 2048 data points in t_2. The DPFGSE sequence [82] was used to suppress water signal. Spectra were processed with the Varian software VNMRJ 1.1D (Varian by Agilent Technologies, Italy) and analyzed with the NEASY [43] program that is included in the CARA software package (http://www.nmr.ch/). NMR analysis was performed for both free peptides by dissolving ~1 mg of peptide in a solution volume equal to 600 µL. NMR analysis was performed in H_2O containing 10% v/v D_2O (99.8% d, Armar Scientific, Switzerland). NMR studies of amphiphilic peptides were not possible because of their high propensity to aggregate.

2.5.8. Molecular Modeling and Dynamics simulations

PeptideK and PeptideE have been built by using the Builder module in InsightII and then subjected to molecular dynamics (MD) simulations performed with the GROMACS software package (v3.3.1) [83]. In detail, each peptide was put in a cubic box filled with SPC216 water molecules, and GROMOS43a1 was selected as force field because it is commonly used as force field for MD simulations on peptides and our group has already used it in other previous papers in which we reported some conformational studies on different peptides where the results obtained from MD simulations resulted in excellent agreement with CD and/or NMR studies [66; 51; 59]. In order to optimize the systems, the peptides were previously subjected to energy minimization and position restraints cycles. The simulations were carried out with periodic boundary conditions by adding chloride ions so that the net electrostatic charge of the system was zero. The bond lengths were constrained by the linear constraint solver algorithm. Particle mesh Ewald algorithm was used for the electrostatic interactions with a cutoff of 0.9 nm, according to our recent papers [66; 51; 72]. All simulations were run for 20 ns at neutral pH and room temperature (300 K) coupling to the system an external bath. GROMACS routines RMSD and RMSF, gyration radius, number of H-bonds and secondary structure evolution) were utilized to check the trajectories and the quality of the simulations. Additionally, on the basis of the distances between structures, like RMSD values, we founded a set of clusters that reflects the range of conformations accessible and the relative weight of each of these by using a clustering algorithm implemented in GROMACS as reported also in our recent papers [66; 59]. The presence of putative H-bonds between the residues and with water molecules was evaluated by Hbplus [84].

2.5.8. Functional Characterization on CXCR4 Receptor

2.5.8.1. Binding assay

PeptideK and PeptideE binding to CXCR4 was evaluated as previously described. [85] CCRF–CEM cells ($2.5x10^5$) were pre-incubated with antagonist peptides (10 μM) in binding buffer (PBS 1x plus 0.2% BSA and 0.1% NaN_3) for 1h at 37°C, 5% CO_2 and then labeled for 45 minutes with anti-CXCR4 PE-antibody (FAB170P, clone 12G5, R&D Systems, Minneapolis, MN, USA). Cells were washed in PBS and analyzed by FACS Canto II cytofluorimeter (Becton Dickinson Immunocytometry Systems, Mountain View, CA, USA).

2.5.8.2. Migration Assay

CCRF-CEM cells migration was assayed in 24-well Transwell chambers (Corning Inc., Corning, NY) using inserts with an 8-μm pore membrane. Membranes were precoated with collagen (human collagen type I/III) and fibronectin (20.0 mg/mL each). CCRF-CEM cells were placed in the upper chamber (1x105 cells/well) in RPMI containing 1% BSA (migration media). Cells were pre-incubated for 45 min with CXCR4 antagonist and allowed to migrate toward 100ng/ml CXCL12 in the lower chamber. After 16 h incubation, migrated cells were collected from the lower chamber and counted. The migration index was defined as the ratio between migrating cells in the experimental group and migrated cells in the control group.

2.5.8.3. cAMP Assay

We have conducted c-AMP assay by using CCRF-CEM cells (1×10^6) that were incubated in the presence of four peptides or Plerixafor (known as AMD3100: a

CXCR4 antagonist) that has provided evidence of concept for inhibition of the CXCR4 pathway [56] at two different concentrations (1 and 10 mM) in combination with forskolin (F) 1 mM for 20 min, followed by stimulation with CXCL12 (100 ng mL^{-1}) for 10 min. Controls include cells stimulated with CXCL12 and forskolin or forskolin alone in the absence of anti-CXCR4 inhibitors. Cells have been harvested and lysed with 0.1 M HCl and cAMP levels have been assayed by a direct competitive enzyme immunoassay (BioVision Incorporated) according to manufacture instructions.

2.6. Appendix

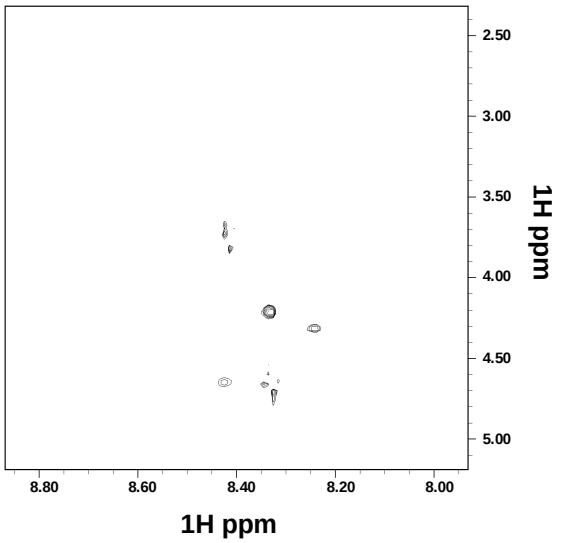


***Figure A1**. 2D [^{1}H, ^{1}H] NOESY 300 spectrum of PeptideE in H_2O.*

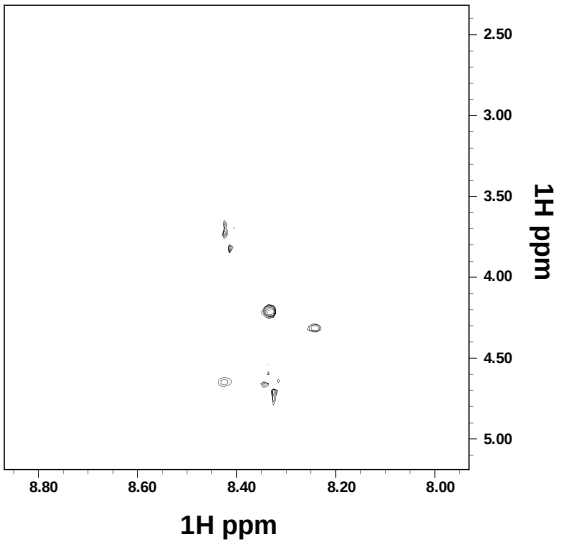


***Figure A2**. 2D [^{1}H, ^{1}H] NOESY 300 spectrum of PeptideK in H_2O.*

***Table A1**. Chemical shift for PeptideE (±0.01) H_2O evaluated at 298 K.*

Residue	H_N	Hα	Hβ	Hγ	Others
♦**1Y**		4.23	3.13		Hδ 7.16 Hε 6.88
2G	8.56	3.98 3.82			
3E	8.46	4.33	2.08 1.97	2.42	
4C	8.47	4.77	2.94 2.87		
5P		4.45	2.30	2.03 1.96	Hδ 3.77
6C	8.38	4.49	2.94		
7Allyl-E	8.58	4.48	2.22 2.02	2.46	Allyl group 5.96, 5.36, 5.30, 4.75

***Table A2**. Chemical shift for PeptideK (±0.01) in in H_2O evaluated at 298 K.*

Residue	H_N	Hα	Hβ	Hγ	Others
♦**1Y**		4.24	3.13		Hδ 7.16, 7.11 Hε 6.88, 6.84
2G	8.57	3.98 3.82			
3E	8.50	4.32	2.04 1.96	2.39	
4C	8.53	4.78	2.91		
5P		4.45	2.30	2.02 1.95	Hδ 3.78
6C	8.39	4.49	2.93		
7Allyl-K	8.61	4.42	1.94 1.80	1.45	Hε 2.99 Hδ 1.68 Allyl group 5.95, 5.34, 5.28, 4.67

♦*The one letter amino acid code, preceded by sequence number, is used to indicate peptide residues.*

***Table A3**. Chemical shift deviations from random coil values (CSD) for PeptideE (D.S. Wishart et al., 1991). All residues are reported by one-letter code. The CSD value for Allyl-E is not shown.*

Residue	$H\alpha_{obs}$	$H\alpha_{randomcoil}$	CSD ($H\alpha_{obs}$-$H\alpha_{randomcoil}$)
♦**1Y**	4.23	4.43	-0.2
2G	3.98 3.82 (3.90)	4.11 3.65 (3.88)	0.02
3E	4.33	4.24	0.09
4C	4.77	4.52	0.25
5P	4.45	4.33	0.12
6C	4.49	4.52	-0.03
7Allyl-E	4.48		

♦*The one letter amino acid code, preceded by sequence number, is used to indicate peptide residues.*

***Table A4**. Chemical shift deviations from random coil values (CSD) for PeptideK (D.S. Wishart et al., 1991). All residues are reported by one-letter code. The CSD value for Allyl-K is not shown.*

Residue	$H\alpha_{obs}$	$H\alpha_{randomcoil}$	CSD ($H\alpha_{obs}$-$H\alpha_{randomcoil}$)
♦**1Y**	4.24	4.43	-0.19
2G	3.98 3.82 (3.90)	4.11 3.65 (3.88)	0.02
3E	4.32	4.24	0.08
4C	4.78	4.52	0.26
5P	4.45	4.33	0.12
6C	4.49	4.52	-0.03
7Allyl-K	4.42		

♦*The one letter amino acid code, preceded by sequence number, is used to indicate peptide residues.*

REFERENCES

1. Paul S.M. *et al*., Nat. Rev. Drug Discov., 2010, 9(3), 203-214.
2. Gonzalez E. and McGraw T. E., Cell Cycle, 2009, 8(16), 2502-2508.
3. Franke T. F., Oncogene, 2008, 27, 6473-6488.
4. Linnerth-Petrik N. M. *et al*., Plos One, 2014, 9(4), e94595.
5. Wu-I W. *et al*., Plos One, 2010, 5(9), e12913.
6. Ropars V. *et al*., Open Spectrosc. J., 2009, 3, 65-76.
7. Yu Y. *et al*., Plos One, 2015, 10(10), e0140479.
8. Berndt N. *et al*., Cell. Death. and Differ., 2010, 17, 1795-1804.
9. Hiromura M. *et al*., J. Biol. Chem., 2004, 279, 53407-53418.
10. Rouse M. B. *et al*., Bioorg. Med. Chem. Lett., 2009, 19, 1508-15011.
11. Barnett S. F. *et al*., Biochem. J., 2005, 385, 399–408.
12. Lindsley C. W., Bioorg. Med. Chem. Lett., 2005,15, 761–764.
13. Desplat V.*et al*., J. Enzym.Inhib. Med. Chem., 2008, 23, 648-658.
14. Desplat V.*et al*., J. Enzym.Inhib. Med. Chem., 2010, 25, 204-215.
15. Desplat V.*et al*., J. Enzym.Inhib. Med. Chem., 2011, 26, 657-667.
16. Jensen H. A. *et al*., Cell. Signal., 2015, 27(8), 1666-1675.
17. Grosdidier A. *et al*., Nucleic Acids Res., 2011, 39(Web Server issue), W270-7.
18. Grosdidier A. *et al*., J. Comput. Chem., 2011, 32(10), 2149-2159.
19. Milburn C. C. *et al*., Biochem. J., 2003, 375, 531–538.
20. Kamijo S. *et al*., J. Am. Chem. Soc.,2005, 127(25), 9260-9266.
21. Gulevich A.V. *et al*., Chem. Rev., 2010, 110, 5235-5331.
22. Liu J. *et al*., Angew. Chem. Int. Edit., 2013, 52, 6953 –6957.
23. Gao M. *et al*., Angew. Chem. Int. Edit., 2013, 52, 6958 –6961.
24. Lichtman M. A., The Oncologist, 2008, 13, 126-138.
25. Campiani G., Butini S., Fattorusso C., Trotta F., Franceschina S., De Angelis M., Nielsen K. S., 2006, WO2006072608.
26. Suzuki M. *et al*., J. Org. Chem., 1974, 39(13), 1980.
27. Guillon J. *et al*., Pharm. Pharm. Comm., 1998, 4, 33-38.
28. Guillon J. *et al*., J. Enzym. Inhib. Med. Chem., 2007, 22(5), 541-549.
29. Ngwerume S., Camp J. E., J. Org. Chem., 2010, 75(18), 6271-6274.
30. Trofimov B. A., Adv. Heterocycl. Chem., 1990, 51, 177-301.
31. Wang Y. F. *et al*., Org. Lett., 2008, 10(21), 5019-5022.
32. Ronga L. *et al*., Eur. J. Med. Chem., 2014, 81, 378-393.
33. Milne J., Normington K. D., Milburn M., 2006, WO2006094210.
34. Sanaeishoar T. *et al*., Comb. Chem. High T. Scr., 2014, 17, 157-161.
35. Yavari I. *et al*., Synlett, 2009, 12, 1921-1922.
36. Sadjadi S. *et al*., Ultrason. Sonochem., 2010, 17, 764-767
37. Bickley J. F. *et al*., Arkivoc, 2002, (vi), 192-204, Gilchrist T. L. & Mendonça R., Synlett, 2000, 12, 1843-1845.
38. Griesinger C. *et al*., J. Am. Chem. Soc., 1988, 110: 7870-7872.
39. Bax A. and Davis D. G., J. Magn. Reson., 1985, 63, 207-213.
40. Piantini U. *et al*., J. Am. Chem. Soc., 1982, 104, 6800-6801;
41. Bax A. and Summers M. F., J. Am. Chem. Soc., 1986, 108, 2093.
42. Bax A. and Subramanian S., J. Magn. Reson., 1986, 67, 565.
43. Bartels C. *et al*., J. Biomol. NMR, 1995, 6, 1-10.
44. Supplementary X-ray crystallographic data: Cambridge Crystallographic Data Centre, University Chemical Lab, Lensfield Road, Cambridge, CB2 1EW, UK; E-mail: deposit@chemcrys.cam.ac.uk.

45. Sheldrick G. M., 1996, SADABS, University of Göttingen, Germany.
46. Sheldrick G. M., Acta Crystallographica Section A, 2008, 64, Part 1, 112-122.
47. Clark D. E. and Pickett S. D., Drug Discov. Today, 2000, 5, 49-58.
48. Pettersen E. F. *et al*., J. Comp. Chem., 2004, 25(13), 1605-12.
49. Liao Y. X. *et al*., Int. J. Mol. Med., 2013, 32, 1239-1246.
50. Sun X., Cheng G., Hao M., *et al*., Cancer Metastasis Rev., 2010, 29, 709-722.
51. Constantini S. *et al*., J. Pept. Sci., 2014, 20(4), 270-278.
52. Dewan M. Z. *et al*., Biomed. Pharmacother., 2006, 60, 273-276.
53. Hideshima T. *et al*, Mol. Cancer Ther., 2002, 1, 539-544.
54. Ponomaryov T. *et al*., J. Clinical Invest., 2000, 106: 1331-1339.
55. Chinni S. R. *et al*., Prostate, 2006, 66, 32-48.
56. Teicher B. A. and Fricker S. P., Clin. Cancer Res., 2010, 16, 2927–2931.
57. Accardo A. *et al*., Mol. BioSyst., 2013, 9, 1401-1410.
58. Uversky V. N., Int. J. Biochem. Cell B., 2011, 43(8), 1090-1103.
59. Palladino P. *et al*., Chem. Biol. Drug Des., 2012, 80, 254–265.
60. Horuk R., Nat. Rev. Drug Discov., 2009, 8, 23–33.
61. Tamamura H. and Fujii N., Expert Opin. Ther. Targets, 2005, 9(6), 1267–1282.
62. Matsumura Y. & Maeda H., Cancer Res., 1986, 46, 6387–6392.
63. Accardo A. *et al*., Mol. BioSyst., 2010, 6, 878–887.
64. Underfriend, S. & Meienhofer, J., In The Peptides, Hruby, V.J., Ed.; Academic Press: New York, NY, USA, 1985; Volume 7.
65. Wiedemnn C. I. *et al*., Bioinformatics, 2013, 29, 1750–1757.
66. Mittal J. *et al.*, J. Phys. Chem. B, 2013, 117(1), 118−124.
67. Sreerama N. & Woody R.W., Biochem., 1994, 33, 10022–10025.
68. Bhatnagar, R. S. & Gough C. A., Circular Dichroism and the Conformational Analysis of Biomolecules; Plenum Press: New York, NY, USA, 1996.
69. Kumar A. *et al*., Biochem. Bioph. Res. Commun., 1980, 95, 1–6.
70. Wishart D. S. *et al*., J. Mol. Biol., 1991, 222, 311–333.
71. Wüthrich, K. NMR of Proteins and Nucleic Acids; John Wiley & Sons: New York, NY, USA, 1986.
72. Guariniello S. *et al*., Biochim. Biophys. Acta, 2014, 1844(2), 447-456.
73. Raucci R. *et al*., Biochim. Biophys. Acta, 2014, 1844(10), 1868-1880.
74. Pappu R. V. *et al*., Proc. Natl. Acad. Sci. USA, 2013, 110, 13392–13397.
75. Vila, J. A. *et al*., Biophys. J., 2004, 86, 731–742.
76. Tarek, M. and Tobias D. J., Phys. Rev. Lett., 2002, 88, 138101.
77. Crump, M. P. *et al*., EMBO J., 1997, 16, 6996–7007.
78. Accardo A. *et al*., Mol. Bio Syst., 2011, 7, 862–870.
79. Morisco A. *et al*., J. Pept. Sci., 2009, 15, 242–250.
80. Accardo A. *et al*., Pept. Sci. Biopolym., 2007, 88, 115–121.
81. Schmitt L. and Dietrich C., J. Am. Chem. Soc., 1994, 116, 8485–8491.
82. Dalvit C., J. Biomol. NMR, 1998, 11, 437-444.
83. Van Der Spoel D. *et al*., J. Comput. Chem., 2005, 26, 1701–1718.
84. McDonald I. K. and Thornton J. M., J. Mol. Biol., 1994, 238, 777–793.
85. Portella L. *et al*., PLoS One, 2013, 8(9), e74548.

SCIENTIFIC PRODUCTION LIST

LIST OF PUBLICATIONS:

1. Leone M., Mercurio F. A., **Vincenzi M.**, Accardo A., Ringhieri P., Tesauro D., Carrière F. and Rossi F., Conformational disorder in phosphopeptides: solution studies by CD and NMR techniques, *Peptidomics*, 2014, 1, 14-21.
2. **Vincenzi M.**, Costantini S., Scala S., Tesauro D., Accardo A., Leone M., Colonna G., Guillon J., Portella L., Ronga L. and Rossi F., Conformational ensembles explored dynamically from intrinsically disordered functionalized peptides: a new class of potential targets in drug discovery, *Special Issue "In-silico prediction and characterization of intrinsic disorder protein", International Journal of Molecular Sciences,* 2015, 16, 12159-12173.
3. **Vincenzi M.**, Accardo A., Costantini S., Scala S., Portella L., Trotta A., Ronga L., Guillon J., Leone M., Colonna G., Rossi F. and Tesauro D., Intrinsically disordered amphiphilic peptides as potential targets in drug delivery vehicles, *Molecular BioSystems*, 2015, 11, 2925-2932.

4. Rubio S., Clarhaut J., Péraudeau E., **Vincenzi M.**, Soum C., Guillon J., Papot S. and Ronga L., Overcoming cyclodimerization in the synthesis of new antiangiogenic cyclic peptide using solution- instead of solid-phase cyclization, *Biopolymers: Peptide Science*, 2016, in press.

5. **Vincenzi M.**, Desplat V., Lucas R., Moreau S., Savrimoutou S., Pinaud N., Lesbordes J., Peyrilles E., Marchivie M., Routier S., Sonnet P., Rossi F., Ronga L. and Guillon J., Synthesis and evaluation of the cytotoxic activity of novel ethyl 4-[4-(4-substitutedpiperidin-1-yl)]benzyl-phenylpyrrolo[1,2-a]quinoxaline-carboxylate derivatives in myeloid and lymphoid leukemia cell lines, *European Journal of Medicinal Chemistry*, accepted for publication.

6. Guillon J., **Vincenzi M.**, Pinaud N., Ronga L., Rossi F., Moreau S., Desplat V. and Marchivie M., Crystal Structure of Ethyl 4-{4-[(4-(5-fluoro-1*H*-benzimidazol-2-yl)piperidin-1-yl)benzyl]}-1-phenylpyrrolo[1,2-*a*]quinoxaline-2-carboxylate, X-Ray Structure Analysis Online, 2016, to be submitted.

7. **Vincenzi M.**, Desplat V., Lucas R., Moreau S., Savrimoutou S., Pinaud N., Bigat D., Enriquez E., Marchivie M., Routier S., Sonnet P., Rossi F., Ronga L. and Guillon J., Synthesis and antiproliferative effect of new 4-(4-substitutedpiperidin-1-yl)]benzylpyrrolo[1,2-*a*]quinoxaline-carboxylate derivatives on human leukemic cells, *European Journal of Medicinal Chemistry*, 2016, to be submitted.

COMMUNICATIONS:

1. **Vincenzi M.**, Ringhieri P., Mercurio F. A., Leone M., Tesauro D., Rossi F. and Accardo A., Solution conformational features and interfacial properties of an intrinsically disordered peptide coupled to alkyl chains: a new class of peptide amphiphiles, **Oral communication**: *1st International Conference on Peptide Materials for Biomedicine and Nanotechnology (PepMat2013)*, Sorrento (Naples), October 28-31, 2013.
2. **Vincenzi M.**, Ringhieri P., Mercurio F. A., Accardo A., Tesauro D., Ronga L., Rossi F. and Leone M., Peptide conformational disorder: solution studies by CD and NMR, **Oral communication**: *21st Young Research Fellow Meeting of SCT (Société de Chimie Thérapeutique)*, Montpellier, March 24-25, 2014.
3. **Vincenzi M.,** Mercurio F. A., Accardo A., Tesauro D., Guillon J., Ronga L., Leone M. and Rossi F., Solution conformational features of intrinsically disordered phosphopeptides: a new class of potential targets in drug discovery, **Poster Presentation**: *14th Naples Workshop on Bioactive Peptides*, Naples, Italy, June 12-14**,** 2014.
4. **Vincenzi M.**, Desplat V., Moreau S., Lucas R., Savrimoutou S., Pinaud N., Rossi F., Ronga L. and Guillon J., Conception, synthèse et évaluation de nouveaux composés hétérocycliques, inhibiteurs potentiels de la protéine kinase Akt, **Poster Presentation**: *Journée Scientifique de la Fédération de Recherche Technologies pour la Santé*, Bordeaux June 24, 2014.
5. **Vincenzi M.**, Desplat V., Moreau S., Savrimoutou S., Lucas R., Rossi F., Ronga L. and Guillon J., Synthesis and Evaluation of the Antiproliferative Activity of Novel Isoindolo[2,1-*a*]quinoxaline and Indolo[1,2-*a*]quinoxaline Derivatives, **Poster Presentation**: *22nd Young Research Fellow Meeting of SCT (Société de Chimie Thérapeutique)*, Romainville, February 4-6, 2015.
6. **Vincenzi M.**, Leone M., Ronga L., Guillon J., Accardo A., Tesauro D., Costantini S. and Rossi F., Could ”disordered” amphiphilic peptides be effective in the modulation of specific bioactive conformations? Preliminary evidences and perspectives. **Oral communication**: *1st NGP-net Symposium on Non-Globular Proteins*, Porto, Portugal, October 6-9, 2015.

Attended Courses and School:

1. 1st Peptide & Protein Chemistry & Biology Training Workshop for Postdoctoral fellows & PhD students, Gennevilliers, October 8-10, 2014
2. Cour de Chimie imagerie et médicaments, Bordeaux, Université de Bordeaux, Septembre 2014-November 2014 (30 Hours)
3. Advanced NMR School organized by GIDRM, Torino, September 1-5 2014
4. Basic NMR School organized by GIDRM, Torino, September 23-27 2013
5. Training Course “Second Generation Bioethanol”, Naples, October 3 18 2013
6. 20/06/2013- Naples (Italy) “From ESTs to RNA-Seq”
7. 27/05/2013- Naples (Italy) “Dynamic processes in microbial metabolism”

As part of a joint program between Italy and France (“co-tutelle de thèse Italie-France”), this PhD work has been carried out at the “Dipartimento di Farmacia”, University “Federico II” of Napoli (Italy), where I spent half of my PhD working time, under the direction of Prof. Filomena Rossi and Dr. Marilisa Leone, and at the “Collège Sciences de la Santé”, Université de Bordeaux (France), where I spent the second half (March 2014-July 2015), under the direction of Prof. Jean Guillon and Dr. Luisa Ronga. My mobility between Italy and France was sponsorized by the “Università italo-francese” of Torino (Italy) through the “Bando VINCI 2014”, project C2-1, 2014/38723, cap. 6.01.1810 UIF.

DOI: 10.2478/ped-2014-0001 — Peptidomics 2014; 14–21 DE GRUYTER OPEN

Research Article **Open Access**

Marilisa Leone#, Flavia Anna Mercurio#, Marian Vincenzi, Antonella Accardo, Paola Ringhieri, Diego Tesauro, Frédéric Carrière, Filomena Rossi*

Conformational disorder in phosphopeptides: solution studies by CD and NMR techniques

Abstract: In the last few years intrinsically disordered proteins (IDPs) have received great attention from the scientific community as they participate in several important biological processes and diseases. The intrinsic disorder and flexibility of IDPs grant them a number of advantages with respect to ordered proteins, such as conformational plasticity to bind several targets, a large interaction surface, involvement in high specificity/low affinity interactions, enhanced binding kinetics. It is assumed that post-translational modifications such as phosphorylation can stimulate structural rearrangement in IDPs and facilitate their binding to partners. To better understand at a structural level the multifaceted mechanisms that govern molecular recognition processes involving IDPs, we designed, synthesized by solid phase methods, and structurally characterized unstructured peptides. These molecules contain a putative disordered module, flanked at either the N- or C-terminal ends by a different phosphorylated amino acid (serine or threonine) to mimick the effects of phosphorylation. The absence of an ordered state in the designed peptides was proved experimentally by CD and NMR conformational studies that were carried out under different solution conditions.

Keywords: intrinsically disordered peptides, peptide design, Solid Phase Peptide Synthesis (SPPS), Circular Dichroism (CD), Nuclear Magnetic Resonance Spectroscopy (NMR).

***Corresponding author Filomena Rossi:** filrossi@unina.it; phone: +39 0812536682; fax: +39 0812536642
Marilisa Leone, Flavia Anna Mercurio, Antonella Accardo, Diego Tesauro and Filomena Rossi: Institute of Biostructures and Bioimaging (IBB-CNR), Via Mezzocannone 16, 80134 Naples, Italy
Marilisa Leone, Antonella Accardo, Paola Ringhieri, Diego Tesauro, Filomena Rossi: Centro Interuniversitario di Ricerca sui Peptidi Bioattivi (CIRPEB), Via Mezzocannone 16, 80134 Naples, Italy
Marian Vincenzi, Antonella Accardo, Paola Ringhieri, Diego Tesauro, Filomena Rossi: University of Naples "Federico II", Department of Pharmacy, Via Mezzocannone 16, 80134 Naples, Italy
Frédéric Carrière: CNRS - Aix-Marseille Université - Enzymologie Interfaciale et Physiologie de la Lipolyse - UMR 7282, 31 chemin Joseph Aiguier, 13402 Marseille cedex 20, France
#These authors contributed equally to the work

1 Introduction

Currently, intrinsically disordered proteins (IDPs) [1,2] receive great attention from the scientific community. IDPs are indeed involved in many important physiological processes related to signaling and regulation of transcription [3], moreover, they take part in several human diseases [4,5].

IDPs are natively disordered proteins that lack in whole or in part a single organized three dimensional structure [1,2]. Recent studies indicate that in eukaryotes more than a third of the proteins are provided with intrinsically disordered regions (IDRs) that are made up of at least thirty residues [2]. It is thought that disorder is actually encoded in the primary sequence and that IDPs are generally enriched in certain types of amino-acids [2,6]. Detailed studies relying on comparison of natively folded and natively unfolded proteins demonstrate that IDPs have generally a low content of "order-promoting" amino acids such as Ile, Leu, Val, Trp, Tyr, Phe, Cys, Asn, whereas, they are considerably enriched in "disorder-promoting" residues (i.e., Ala, Arg, Gly, Gln, Ser, Glu, Lys and Pro) [1]. Intrinsically disordered regions (IDRs) have in common several characteristics such as flexibility, β-sheet propensity, low average hydrophobicity and high net charge; these features are hallmarks of IDPs and have been used to develop specific predictors of disorder [1].

IDPs can be considered "promiscuous", as they can bind through several interactions, different targets [7]. Binding of IDPs to partners may result in formation of highly dynamic complexes [8-10] or be accompanied by a folding process with a generation of an ordered structure. This latter process allows IDPs to gain a particular biological function.

Interestingly, it has been reported that several peptides, such as the CCK8, to interact selectively with their receptors require an initial disordered conformation [11].

For the coupled "binding-folding" of IDPs a "fly-casting" mechanism has been proposed that is basically responsible for enhanced interaction kinetics [12,13]. In contrast to ordered proteins, the large chain flexibility in IDPs supplies a bigger capture radii; according to the "fly-casting" model, a flexible region, that represents the unfolded state, will first interact partially and weakly with its target relatively far from the binding site and then, will fold while approaching to it [13].

IDPs often contain post-translational modifications (PTM), such as phosphorylation, that can stimulate structural rearrangements. It is hypothesized that the location of PTM sites in proximity of a disordered segment may facilitate the binding of the enzyme catalyzing the PTM in several targets [1,5].

To get better understanding of the complex machineries that govern molecular recognition processes involving IDPs, we have undertaken studies of predicted unstructured peptides (See Table 1). In major detail, peptide sequences enriched in disorder promoting amino acids [1] flanked at either the N- or C-terminal end by a phosphorylated residue, were designed and synthesized by solid phase synthesis methods. In addition, the conformational behavior of the following unphosphorylated peptide sequence (IDP), NH_2-AQIREASSPSLQVDNQSDQT*g*-$CONH_2$ (*g*=Allyl-glycine), was also investigated. The absence of one single ordered conformation in these peptides was established experimentally by CD and NMR analyses that were conducted under different solution conditions (i.e.: aqueous buffer, water/trifluoroethanol (TFE) mixtures, dimethylsulfoxide (DMSO)).

These peptides may represent new disordered scaffolds which, by preserving most of the advantages of IDPs, could eventually take part in several molecular recognition processes and exert biological actions.

2 Experimental Procedures

2.1 Disorder prediction

A peptide sequence with a high disorder propensity, i.e. a region enriched in "disorder promoting residues" (A, R, Q, S, P and E) [1] was designed and checked for disorder prediction using the MeDor metaserver (http://www.vazymolo.org/MeDor/index.html) [14].

2.2 Chemicals

All solvents were reagent grade. N^{α}-Fmoc-protected amino acids and activating agents were purchased from Inbios (Pozzuoli, Italy), N^{α}-Fmoc-protected-D-phospho-Threonine (*pt*) and N^{α}-Fmoc-protected-L-phospho-Serine (*pS*) from AnaSpec, Inc. (Fremont, CA). Resin for peptides synthesis was purchased from Novabiochem (Läufelfingen, Switzerland). All other chemicals were commercially available from Sigma-Aldrich, Fluka (Buchs, Switzerland) or LabScan (Stillorgan, Dublin, Ireland) and were used as received, unless otherwise stated. HPLC chemicals were purchased from Lab-Scan (Dublin, Ireland).

2.3 Peptide synthesis

IDP-phosphopeptides and IDP-allyl-glycine (see Table 1) were synthesized in batch by using standard solid-phase 9-fluorenylmethoxycarbonyl (Fmoc) procedures [15] on Rink-amide 4-methylbenzhydrylamine (MBHA) resin (0.65 mmol g^{-1}, 0.020 mmol scale). Peptide elongation was achieved by sequential addition of Fmoc-AA-OH using 1-hydroxybenzotriazole (HOBt)/ O-benzotriazole-tetramethyl-uronium-hexafluoro-phosphate (HBTU)/ diisopropylethylamine (DIPEA) (1/1/2 v/v/v) as coupling reagents, in dimethylformamide (DMF) in preactivation mode. All couplings were performed for 30 minutes, by using an excess of 5 equivalents for the single amino acid derivative. Fmoc removal was achieved by piperidine/DMF (3:7, v/v) treatment for 10 min. Peptides deprotection and cleavage from the solid support were achieved by treatment with a trifluoroacetic acid (TFA)/ triisopropylsilane (TIS)/water (95/2.5/2.5, v/v/v) mixture for 90 min at room temperature. The crude peptides were precipitated at 0 °C with ethyl ether, dissolved in a water/ acetonitrile (1:1, v/v) mixture and lyophilized. Crude products were purified by RP-HPLC chromatography. Preparative RP-HPLCs were carried out on a LC8 Shimadzu HPLC system (Shimadzu Corporation, Kyoto, Japan)

Table 1. Peptide sequences and their molecular weight.

Peptide	Sequences	M_W (Da)
IDP1	NH_2-AQIREASSPSLQVDNQSDQ*pt*-$CONH_2$	2254.2
IDP2	NH_2-*pt*AQIREASSPSLQVDNQSDQT-$CONH_2$	2355.4
IDP3	NH_2-*pS*AQIREASSPSLQVDNQSDQT-$CONH_2$	2341.4
IDP	NH_2-AQIREASSPSLQVDNQSDQT*g*-$CONH_2$	2270.2

pt=D-phospho-Threonine; pS=L-phospho-Serine, g=Allyl-glycine

equipped with an UV lambda-Max Model 481 detector, by using a Phenomenex (Torrance, CA) -C18 Jupiter column (90 Å; 10 mm) and by using as elution solvents 0.1% TFA in water (a) and 0.1% TFA in acetonitrile (b) from 5 to 70% over 30 min (flow rate: 20 mL/min). Peptides purity and integrity were assessed by analytical LC–MS analysis with a Finnigan Surveyor MSQ single quadrupole electrospray ionization spectrometer (Finnigan/ Thermo Electron Corporation San Jose, CA), by eluting with 0.1% TFA in water (a) and 0.1% TFA in acetonitrile (b) from 5 to 70% over 15 min (0.8 mL/min). Characterization was conducted under standard conditions of peptide analysis. The expected and experimental mass of our peptides are the following: M.W. IDP1=2254.2 Da and $[M+2H^+]/2$=1128.1 m/z; M.W. IDP2=2355.4 Da and $[M+2H^+]/2$=1178.7 m/z; M.W. IDP3=2341.4 Da and $[M+2H^+]/2$=1171.7 m/z; M.W. IDP= 2270.2 Da and $[M+2H^+]/2$= 1135.1 m/z.

2.4 Circular dichroism measurements

Far-UV CD spectra of IDP1 were recorded from 190 to 260 nm on a Jasco J-810 spectropolarimeter equipped with a NesLab RTE111 thermal controller unit using a 0.1 cm quartz cell at 25 °C. Circular dichroism measurements were carried out on peptide solutions at $1{\cdot}10^{-4}$ M concentration. Solutions were prepared by promptly dissolving lyophilized pure peptide powder in net H_2O at a final pH ~ 5 (the acidic pH was due to the residual TFA used during peptide cleavage and purification). CD spectra were also recorded for an IDP1 sample dissolved in a mixture of water/trifluoroethanol (TFE) (50/50) and of water/TFE (25/75).

Other experimental settings were: scan speed, 10 nm min^{-1}; sensitivity, 50 mdeg; time constant, 4 s; bandwidth, 2 nm. Each spectrum was obtained by averaging three scans, subtracting contributions from other species in solution and converting the signal to mean residue ellipticity in units of deg cm^2 $dmol^{-1}$ res^{-1}.

2.5 Nmr measurements and solution structure calculations

NMR analysis was performed for IDP, IDP1, IDP2 and IDP3. NMR samples were prepared by dissolving about 0.5 mg of each peptide in a solution volume equal to 600 µL. NMR analysis of IDP1 was performed in H_2O containing 10% v/v D_2O (99.8% d, Armar Scientific, Switzerland) and in a mixture H_2O/trifluoroethanol-d3 (98% d, Armar Chemicals, Switzerland) 17/83 v/v. Solution conformational studies of IDP were conducted in H_2O/trifluoroethanol-d3 (98% d, Armar Chemicals, Switzerland) 17/83 v/v. NMR characterizations of IDP2 and IDP3 were carried out in DMSO-d6 (99.9% d, Armar Chemicals, Switzerland).

NMR spectra were recorded at 25 °C on a Varian Unity Inova 600 MHz spectrometer provided with a cold probe. The process of proton resonance assignments (See Supplemental Tables S1, S2, S3 and S4) was carried out with a canonical protocol [16] based on analysis of the following two dimensional [^{1}H, ^{1}H] spectra: TOCSY (Total Correlation Spectroscopy) (70 ms mixing time) [17], DQFCOSY (Double Quantum Filter Correlation Spectroscopy) [18], NOESY (Nuclear Overhauser Enhancement Spectroscopy) [19] (200 and 300 ms mixing times). Chemical shifts were referenced with respect to the TSP (Trimethylsilyl-3-propionic acid sodium salt-d4, 99% d, Armar Scientific, Switzerland) signal at 0.0 ppm.

1D spectra were acquired with a relaxation delay of 1s and 32-128 scans. 2D experiments were generally acquired with 32-64 scans, 128-256 FIDs in t_1, 1024 or 2048 data points in t_2.

The DPFGSE (Double Pulsed Field Gradient Selective Echo) sequence [20] was used to suppress water signal. Spectra were processed with the Varian software VNMRJ 1.1D (Varian by Agilent Technologies, Italy) and analyzed with the NEASY [21] program that is included in the CARA (Computer Aided Resonance Assignment) software package (http://www.nmr.ch/).

Peptide structure calculations were carried out with the software CYANA (version 2.1) [22]. A D-phospho-Threonine and L-phospho-Serine were introduced in the CYANA standard residue library [23]. Distance constraints for structure calculations were gained from NOESY experiments (300 ms mixing time). The GRIDSEARCH module of CYANA was used to generate angular constraints. Calculations started from 100 random conformers; the 20 conformers with the lowest CYANA target functions were finally checked with the program MOLMOL [24] and iCing (http://proteins.dyndns.org/cing/iCing.html) [25].

3 Results and Discussion

Four intrinsically-disordered peptides (IDP, IDP1, IDP2 and IDP3; see Table 1) were designed and analyzed. Their primary sequences include a predicted disordered region (-AQIREASSPSLQVDNQSDQ-) (See also Material and Methods for details) and a phosphorylated amino acid at C-terminus (IDP1 with phospho-Threonine) or N-terminus (IDP2 with phospho-Threonine and IDP3 with phospho-Serine, respectively) or an Allyl-glycine at the C-terminus (IDP).

The high disorder propensity of the common -AQIREASSPSLQVDNQSDQ- peptide core was evaluated

with the MeDor server [14] (Supplemental Figure S1, upper panel) while its tendency towards β-aggregation was investigated with the software TANGO [26] which indicated that its sequence is not inclined to form β-aggregates (Supplemental Figure S1, lower panel).

Conformational features of IDP, IDP1, IDP2 and IDP3 were studied in solution by CD and NMR techniques.

3.1 Circular Dichroism

CD spectra of IDP1 in water and in the water/TFE mixtures (50/50 and 25/75 v/v) were recorded between 190 and 260 nm. IDP2 and IDP3 peptide sequences were not soluble under the same experimental conditions. The shape of CD spectrum of IDP1 in water, with a negative band between 197 and 201 nm, suggests an unordered structure (see Figure 1A). The addition of TFE induces a conformational rearrangement from random coil to β-structure. This rearrangement can be better appreciated by difference spectra (Figure 1B), which show a pronounced negative maximum around 220 nm and a positive maximum around 200 nm, which are characteristic of β-sheet conformation. However, the low structural content gained by the peptide, even in TFE, supports its high flexibility. Similar CD spectra were also detected for the unphosphorylated IDP peptide at different water/TFE ratios (data not shown).

3.2 NMR studies

Detailed NMR studies of IDP, IDP1, IDP2 and IDP3, including acquisition and analysis of 1D [^{1}H] and 2D [^{1}H, ^{1}H] spectra, were conducted at 25°C.

We first investigated the conformational preferences of IDP1 in H_2O/D_2O (90/10 v/v). Under these experimental conditions, the poor spectral dispersion of the 1D [^{1}H] spectrum (Figure 2A) and the almost complete absence of signal in the 2D [^{1}H, ^{1}H] NOESY experiment indicated that IDP1 was very flexible and disordered (Supplemental Figure S2A).

Next we treated IDP1 with TFE that is a useful co-solvent to investigate the inherent conformational preferences of proteins and peptides [27] and has already been used in IDPs studies [28-30]. Since CD data indicated a gain of structure in IDP1 only at high TFE concentration, we conducted NMR characterization of the peptide in presence of 83% TFE.

A gain of ordered structure induced by TFE can already be appreciated in the H_N region of the 1D proton experiment (Figure 2A) where an improvement of the spectral dispersion is evident. Nevertheless, many cross-peaks appear in the 2D [^{1}H, ^{1}H] NOESY spectrum as well (Figure 2C) possibly indicating a decrease of flexibility in the peptide. The high quality of the spectra, recorded in presence of TFE, allowed us to carry out a complete structural characterization and obtain proton resonance assignments (Figure 2 and Supplemental Table S1). We evaluated chemical shift deviations (CSD) of Hα protons from random coil values (Supplemental Figure S3) and noticed that they did not point to any specific secondary structure element. In fact, most of the CSD were low (i.e. with absolute values lower that 0.1) and positive (positive deviations are generally indicative of extended conformations) [31]. Afterwards, we analyzed short and medium range NOEs (Figures 3A and S4). For most of IDP1

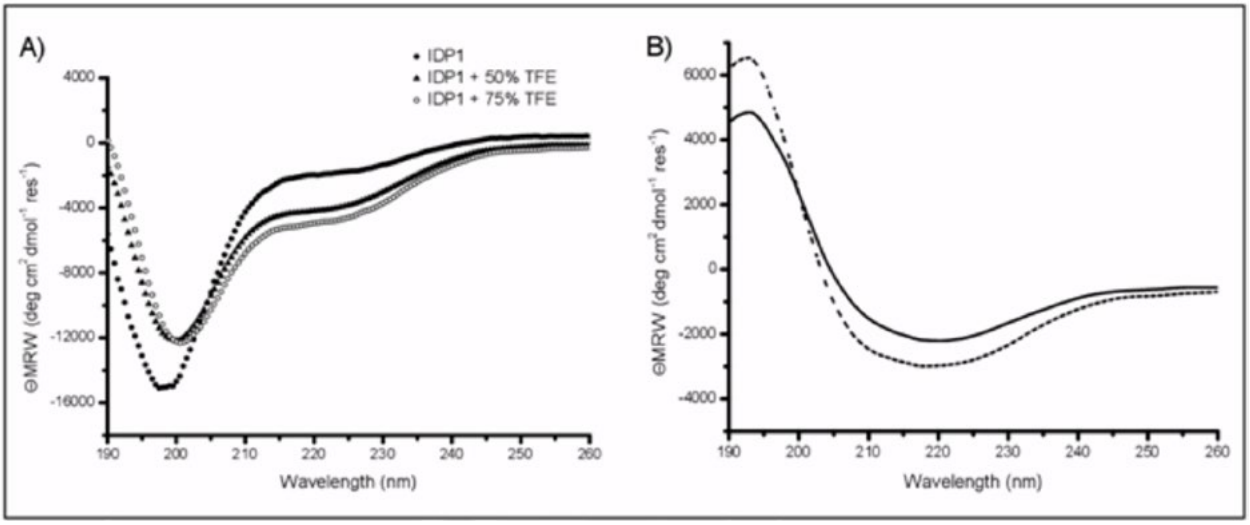


Figure 1. Far UV CD spectra of: A) IDP1 at 0, 50 and 75% of TFE and of B) IDP1 at 50% (solid line) and 75% (dash line) of TFE after subtraction of IDP1 at 0% TFE.

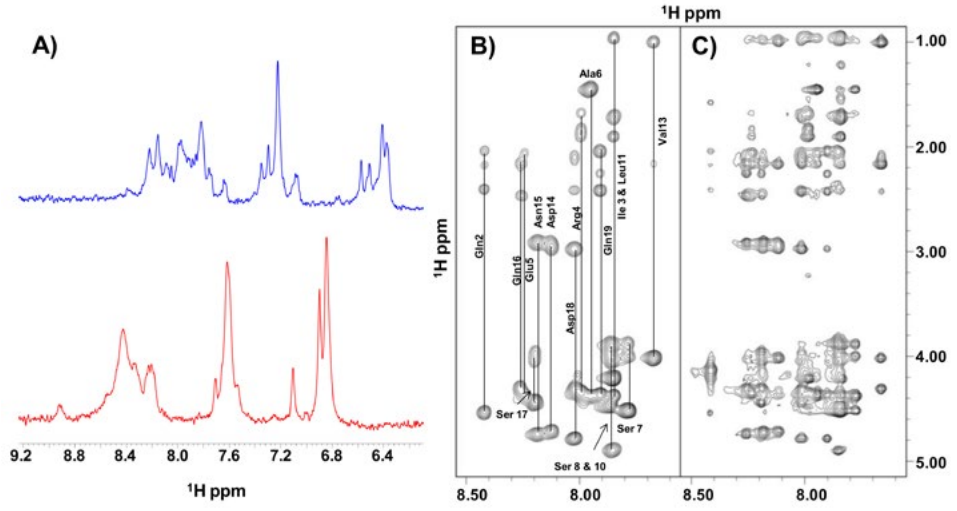


Figure 2. (A) Overlay of 1D proton spectra of IDP1 recorded in H_2O/D_2O (90/10 v/v) (red) and in H_2O/TFE-d3 (17/83 v/v) (blue). (B) 2D [^{1}H, ^{1}H] TOCSY and (C) 2D [^{1}H, ^{1}H] NOESY 300 spectra of IDP1 acquired in H_2O/TFE-d3 (17/83 v/v); the H_N/aliphatic protons correlation regions of both spectra are shown; spin system assignments are indicated in (B).

primary sequence sequential NOE contacts of the type $H\alpha_i$-HN_{i+1} (indicated as $d_{\alpha N(i, i+1)}$ in Fig. 3A), $H\beta_i$-HN_{i+1} ($d_{\beta N(i, i+1)}$ in Fig. 3A) and HN_i-HN_{i+1} ($d_{NN(i, i+1)}$ in Fig. 3A) are predominant and not clearly indicative of any ordered conformation [16]. Interestingly, the NOE pattern improves in the segment encompassing residues from Ser10 to Gln16, where medium range contacts typical of helical structures (i.e. $d_{\alpha\beta(i, i+3)}$ and $d_{\alpha N(i, i+3)}$) can be observed [16]. The Pro residue is mainly in trans configuration as shown by the strong NOE cross peak in between the Hα proton of Ser8 and the Hδ protons of Pro9, however, a weaker NOE between the Hα protons of Ser8 and Pro9 is evident as well and points to some proline cis-trans isomerization [16].

Indeed, structural calculations, carried out with the software CYANA [22] (See Table 2 and Figure 3), demonstrate that the peptide is rather flexible and assumes a more ordered pseudo-helical turn only between residues Val13 and Gln16 (Figure 3B,C). This small helical/turn contribution to the overall IDP1 disordered conformation in TFE cannot be appreciated by CD experiments which mainly underline the prevalence of extended structures (Figure 1).

The disorder of the peptide is represented by the high RMSD values measured for the NMR ensemble (Figure 3B and Table 2). In summary, these data, in agreement with CD results, indicate that the propensity of IDP1 to gain an ordered secondary structure is rather low even under strong structuring conditions (i.e. high percentage of TFE in solution). For comparison purpose, we also carried out similar NMR studies in aqueous solution, containing 83% TFE, of the analogue peptide: AQIREASSPSLQVDNQSDQTg (where g= Allyl-glycine) which lacks a phosphorylated residue (Supplemental Figure S5 and Table S2). This peptide exhibits identical conformational behavior as IDP1, as indicated by similar CD spectrum (data not shown) and NMR parameters such as CSD (Supplemental Figure S6) and NOE pattern (Supplemental Figure S7). These data likely demonstrate that the presence of a phosphate group does not dramatically influence the overall conformational disorder of our IDP peptides.

Attempts to carry out structural studies of IDP2 and IDP3 in water and in water/TFE mixtures were hampered by the low solubility of both peptides. Thus, we analyzed their structural features in DMSO, a solvent in which it was possible to record high quality NMR data (Figures 4 and 5) and obtain proton resonance assignments (Supplemental Tables S3 and S4). In DMSO, IDP2 and IDP3 show a very similar conformational behavior. 2D [^{1}H, ^{1}H] NOESY spectra of both peptides (Figure 4B and 5B) contain many cross-peaks and resemble typical experiments recorded for rigid folded species. Detailed analysis of NOE patterns reveals a clear prevalence of sequential contacts, such as $H\alpha_i$-HN_{i+1}, $H\beta_i$-HN_{i+1}, HN_i-HN_{i+1}, (See Figure 4C) that alone are not sufficient to indicate any specific ordered secondary structure element [16]. However, these kinds of NOEs co-exist with several medium range contacts in between side chain protons two residues apart in the sequence, and with sequential Hα-Hα contacts, the latter are concentrated in the peptide region 15-DNQSDQ-20, and are generally more characteristic of beta type structures. Moreover, in

Table 2. Structure statistics for the IDP1 NMR ensemble calculated in H_2O/TFE-d3 (17/83 v/v).

NOE upper distance limits	143
Angle constraints	88
Residual target function, Å²	0.22±0.07
Residual NOE violations	
Number > 0.1 Å#	2
Number	7
Residual angle violations	
Number	0
Atomic pairwise RMSD@, Å	
Backbone atoms (a.a. 2-19)	2.39±0.54
Heavy atoms (a.a. 2-19)	3.57±0.66

Average CYANA violation; @Calculated by iCing

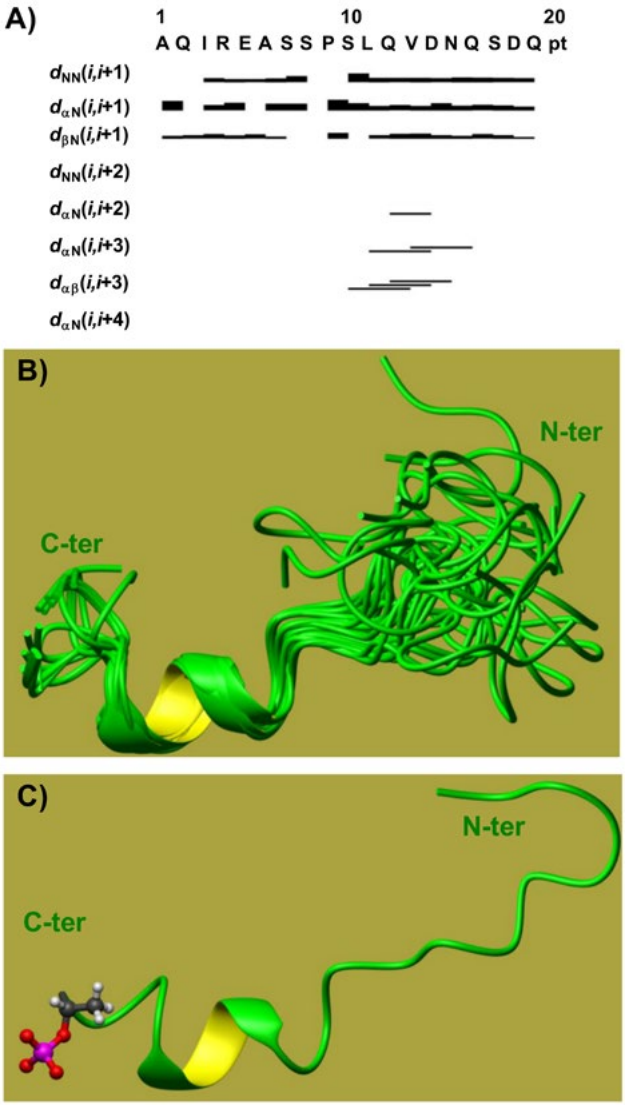


Figure 3. (A) NOE intensity pattern for IDP1 in H_2O/TFE-d3. Only principal short and medium range NOEs are reported. Different residues are specified with the one-letter amino acid code; "$d_{\xi\zeta}$ (b, c)" designates a NOE contact between the H_ξ and H_ζ protons in the b and c residues respectively. (B) Overlay on the backbone atoms (residues 12-18, RMSD=0.70 ± 0.33 Å) of 20 IDP1 structures. (C) Ribbon representation of the best IDP1 CYANA conformer (i.e. the one with the lowest target function). The D-phospho-Threonine side chain is reported in neon representation. The final CYANA calculation included 143 upl (upper distance limits) of which 60 intra-residue, 73 short- and 10 medium-range.

analogy with the IDP1 peptide, the Pro residue in IDP2 and IDP3 contributes to the overall disorder. Indeed the trans configuration is predominant but some cis-trans interconversion is evident as well [16]. Tentative complete structure calculations were carried out and demonstrated the intrinsic conformational disorder of IDP2 and IDP3 (See Figure S8). The presence of several types of NOEs in the spectra of IDP2 and IDP3 may indeed point to the occurrence of either inter-conversion of unfolded and partially ordered states or towards aggregation processes that lower the flexibility of the two peptides in DMSO. It's worth nothing that many intrinsically disordered peptides exhibit high tendency to aggregate, and in fact, under physiological conditions, are able to form beta fibrils ("amyloids") [32,33] and play roles in pathological processes like Alzheimer's and Parkinson's diseases [34-36]. Based only on the low computationally predicted tendency towards β-aggregation of the peptide core -AQIREASSPSLQVDNQSDQ- (Supplemental Figure S1, lower panel), this latest scenario appears unlikely for IDP2 and IDP3. On the other side, we cannot exclude that the addition of a phosphorylated amino acid may change the predicted aggregation propensity and/or the implemented non physiological conditions may favor peptide aggregation.

4 Conclusions

IDPs represent targets in drug discovery for several diseases (for example tumor, Parkinson's disease, Alzheimer's disease, type II diabetes) [5,6,37]. In order to create novel and efficient drug discovery strategies for such proteins, it is essential to fully unveil distinctive characteristics of intrinsically disordered regions. To this aim we designed and synthesized three intrinsically disordered peptides, made up of a predicted disordered segment flanked at either the N- and C-terminal side by a phospho-Serine or a phospho-Threonine (See Table 1). We performed NMR

and CD solution conformational studies. In particular we first analyzed the peptide IDP1, that contains a phospho-Threonine at the C-terminal end. The absence of precise secondary structure elements was proved in water and water/TFE mixtures. The addition of TFE in high percentage only slightly decreases IDP1 conformational flexibility thus stressing out the natively disordered nature of this peptide. The same conformational trend in a solution containing TFE, is exhibited by the analogue unphosphorylated IDP peptide (See Table 1); these data likely indicate that phosphorylation does not affect relevantly the structure of our peptides.

We next analyzed IDP2 and IDP3 whose primary sequences differ only in the first N-terminal residue: either a phospho-Threonine or a phospho-Serine. Solubility issues did not allow us to characterize these two peptides in aqueous buffer, we thus investigated their conformational preferences in DMSO. Under these non physiological conditions, IDP2 and IDP3 exhibited identical conformational behavior, characterized by the absence of one single ordered state and probably by aggregation phenomena.

As concerning the possible applications of the disordered IDP1, IDP2 and IDP3 peptides, we can certainly envision their use to generate novel peptide amphiphiles (PAs) to be implemented in the field of biomedicine and biocompatible materials [11,38-41].

In fact, we have very recently described novel PAs, obtained by coupling an intrinsically disordered short peptide to different alkyl-chains, that are able to self-assemble in supramolecular aggregates (i.e., unilamellar vesicles and micelles) with a consequent reduction of flexibility and a gain of structure with respect to the free peptide [41].

Moreover, we cannot exclude that our IDP1, IDP2 and IDP3 peptides, being provided with characteristic features of intrinsically disordered proteins, may take part in different protein-protein interaction networks and have a particular biological action, but of course deep studies in a cellular environment are needed to prove this scenario.

Acknowledgements: The authors thank Dr. Sonia Longhi for helping with the design of disordered peptide sequences; Rosa Aufiero for preliminary peptide synthesis; Leopoldo Zona and Luca De Luca for technical assistance. F.A.M. thanks Regione Campania P.O.R. Campania FSE 2007/2013 Project STRAIN for the fellowship.

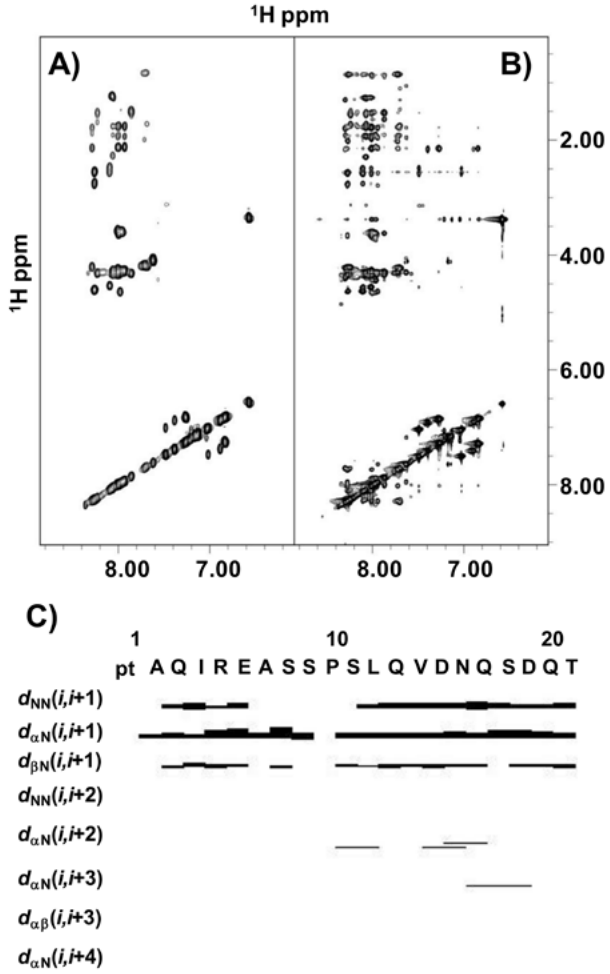


Figure 4. Comparison of (A) 2D [¹H, ¹H] TOCSY and (B) NOESY 300 spectra of IDP2 recorded in DMSO-d6. The H_N correlation regions of the spectra are shown. (C) NOE intensity pattern for IDP2.

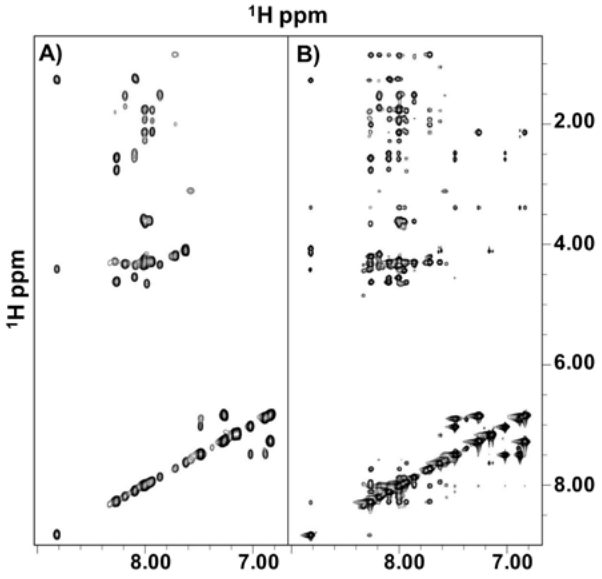


Figure 5. Expansion of a region of (A) 2D [¹H, ¹H] TOCSY and (B) 2D [¹H, ¹H] NOESY 300 of IDP3 where correlations arising from H_N protons can be observed. Spectra were acquired in DMSO-d6.

References

[1] V. N. Uversky, Int J Biochem Cell Biol 2011, 43, 1090-1103.
[2] M. M. Babu, Mol Biosyst 2012, 8, 21.
[3] J. J. Ward, J. S. Sodhi, L. J. McGuffin, B. F. Buxton, D. T. Jones, J Mol Biol 2004, 337, 635-645.
[4] Y. Cheng, T. LeGall, C. J. Oldfield, A. K. Dunker, V. N. Uversky, Biochemistry 2006, 45, 10448-10460.
[5] M. M. Babu, R. van der Lee, N. S. de Groot, J. Gsponer, Curr Opin Struct Biol 2011, 21, 432-440.
[6] V. N. Uversky, Expert Rev Proteomics 2010, 7, 543-564.
[7] G. P. Singh, D. Dash, Proteins 2007, 68, 602-605.
[8] A. B. Sigalov, Adv Exp Med Biol 2008, 640, 268-311.
[9] A. B. Sigalov, Mol Biosyst 2010, 6, 451-461.
[10] A. B. Sigalov, Self Nonself 2010, 1, 89-102.
[11] A. Accardo, A. Morisco, P. Palladino, R. Palumbo, D. Tesauro, G. Morelli, Mol Biosyst 2011, 7, 862-870.
[12] Y. Huang, Z. Liu, J Mol Biol 2009, 393, 1143-1159.
[13] B. A. Shoemaker, J. J. Portman, P. G. Wolynes, Proc Natl Acad Sci U S A 2000, 97, 8868-8873.
[14] P. Lieutaud, B. Canard, S. Longhi, BMC Genomics 2008, 9 Suppl 2, S25.
[15] W. Chang, P. D. White, Fmoc Solid Phase Peptide Synthesis. , Oxford University Press, New York, 2000.
[16] K. Wüthrich, NMR of Proteins and nucleic acids., John Wiley & Sons, New York, 1986.
[17] C. Griesinger, G. Otting, K. Wüthrich, R.R. Ernst, J Am Chem Soc 1988, 110, 7870-7872.
[18] U. Piantini, O.W. Sorensen, R.R. Ernst, J Am Chem Soc 1982, 104, 6800-6801.
[19] A. Kumar, R. R. Ernst, K. Wuthrich, Biochem Biophys Res Commun 1980, 95, 1-6.
[20] C. Dalvit, Journal of Biomolecular NMR 1998, 11, 437-444.
[21] C. Bartels, T. H. Xia, M. Billeter, P. Guntert, K. Wuthrich, J Biomol NMR 1995, 6, 1-10.
[22] T. Herrmann, P. Guntert, K. Wuthrich, J Mol Biol 2002, 319, 209-227.
[23] J. W. Craft, Jr., G. B. Legge, J Biomol NMR 2005, 33, 15-24.
[24] R. Koradi, M. Billeter, K. Wuthrich, J Mol Graph 1996, 14, 51-55, 29-32.
[25] J. F. Doreleijers, A. W. Sousa da Silva, E. Krieger, S. B. Nabuurs, C. A. Spronk, T. J. Stevens, W. F. Vranken, G. Vriend, G. W. Vuister, J Biomol NMR 2012, 54, 267-283.
[26] A. M. Fernandez-Escamilla, F. Rousseau, J. Schymkowitz , L. Serrano. Nat Biotechnol. 2004, 10, 1302-1306.
[27] M. Buck, Q Rev Biophys 1998, 31, 297-355.
[28] K. C. Hite, A. A. Kalashnikova, J. C. Hansen, Protein Sci 2012, 21, 531-538.
[29] B. Maestro, B. Galan, C. Alfonso, G. Rivas, M. A. Prieto, J. M. Sanz, PLoS One 2013, 8, e56904.
[30] V. L. Anderson, T. F. Ramlall, C. C. Rospigliosi, W. W. Webb, D. Eliezer, Proc Natl Acad Sci U S A 2010, 107, 18850-18855.
[31] D. S. Wishart, B. D. Sykes, F. M. Richards, J Mol Biol 1991, 222, 311-333.
[32] S. Ohnishi, K. Takano, Cell Mol Life Sci 2004, 61, 511-524.
[33] F. Rousseau, J. Schymkowitz, L. Serrano, Curr Opin Struct Biol 2006, 16, 118-126.
[34] M. von Bergen, P. Friedhoff, J. Biernat, J. Heberle, E. M. Mandelkow, E. Mandelkow, Proc Natl Acad Sci U S A 2000, 97, 5129-5134.
[35] M. R. Krebs, D. K. Wilkins, E. W. Chung, M. C. Pitkeathly, A. K. Chamberlain, J. Zurdo, C. V. Robinson, C. M. Dobson, J Mol Biol 2000, 300, 541-549.
[36] A. Thompson, A. R. White, C. McLean, C. L. Masters, R. Cappai, C. J. Barrow, J Neurosci Res 2000, 62, 293-301.
[37] J. Wang, Z. Cao, L. Zhao, S. Li, Int J Mol Sci 2011, 12, 3205-3219.
[38] H. Cui, M. J. Webber, S. I. Stupp, Biopolymers 2010, 94, 1-18.
[39] A. Accardo, G. Mangiapia, L. Paduano, G. Morelli, D. Tesauro, J Pept Sci 2013, 19, 190-197.
[40] C. Falciani, J. Brunetti, B. Lelli, A. Accardo, D. Tesauro, G. Morelli, L. Bracci, J Pept Sci 2013, 19, 198-204.
[41] A. Accardo, M. Leone, D. Tesauro, R. Aufiero, A. Benarouche, J. F. Cavalier, S. Longhi, F. Carriere, F. Rossi, Mol Biosyst 2013, 9, 1401-1410.

Int. J. Mol. Sci. **2015**, *16*, 12159-12173; doi:10.3390/ijms160612159

OPEN ACCESS
International Journal of
Molecular Sciences
ISSN 1422-0067
www.mdpi.com/journal/ijms

Article

Conformational Ensembles Explored Dynamically from Disordered Peptides Targeting Chemokine Receptor CXCR4

Marian Vincenzi [1,2,6,†], Susan Costantini [3,†], Stefania Scala [4], Diego Tesauro [1,2], Antonella Accardo [1,2], Marilisa Leone [2], Giovanni Colonna [5], Jean Guillon [6], Luigi Portella [4], Anna Maria Trotta [4], Luisa Ronga [6,*] and Filomena Rossi [1,2,*]

1 Department of Pharmacy University of Naples "Federico II", and CIRPeB, Via Mezzocannone 16, I-80134 Naples, Italy; E-Mails: marian.vincenzi@unina.it (M.V.); diego.tesauro@unina.it (D.T.); antonella.accardo@unina.it (A.A.)

2 Institute of Biostructures and Bioimaging (IBB), National Research Council (CNR), Via Mezzocannone 16, I-80134 Naples, Italy; E-Mail: marilisa.leone@cnr.it

3 Oncology Research Center of Mercogliano (CROM), Istituto Nazionale Tumori "Fondazione G. Pascale", IRCCS, I-80131 Napoli, Italy; E-Mail: s.costantini@istitutotumori.na.it

4 Molecular Immunology and Immuneregulation, Istituto Nazionale Tumori "Fondazione G. Pascale", IRCCS, I-80131 Napoli, Italy; E-Mails: s.scala@istitutotumori.na.it (S.S.); portella@gmail.com (L.P.); amt78@libero.it (A.M.T.)

5 Center of Medical Informatics, AOU, Second University of Naples, I-80138 Napoli, Italy; E-Mail: giovanni.colonna@unina2.it

6 UFR des Sciences Pharmaceutiques—Collège Sciences de la Santé, INSERM U869, Laboratoire ARNA, Université de Bordeaux, 146 rue Léo Saignat, 33076 Bordeaux cedex, France; E-Mail: jean.guillon@u-bordeaux.fr

† These authors contributed equally to this work.

* Authors to whom correspondence should be addressed; E-Mails: luisa.ronga@u-bordeaux.fr (L.R.); filomena.rossi@unina.it (F.R.); Tel.: +33-0-547-4730-4264 (L.R.); +39-081-253-6682 (F.R.); Fax: +39-081-253-6642 (F.R.).

Academic Editor: Lukasz Kurgan

Received: 2 March 2015 / Accepted: 20 May 2015 / Published: 28 May 2015

Abstract: This work reports on the design and the synthesis of two short linear peptides both containing a few amino acids with disorder propensity and an allylic ester group at the

C-terminal end. Their structural properties were firstly analyzed by means of experimental techniques in solution such as CD and NMR methods that highlighted peptide flexibility. These results were further confirmed by MD simulations that demonstrated the ability of the peptides to assume conformational ensembles. They revealed a network of transient and dynamic H-bonds and interactions with water molecules. Binding assays with a well-known drug-target, *i.e.*, the CXCR4 receptor, were also carried out in an attempt to verify their biological function and the possibility to use the assays to develop new specific targets for CXCR4. Moreover, our data indicate that these peptides represent useful tools for molecular recognition processes in which a flexible conformation is required in order to obtain an interaction with a specific target.

Keywords: intrinsically disordered protein (IDP); intrinsically disordered region (IDR); conformational ensemble; MD; NMR; CD; chemokine

1. Introduction

In the last decade, several proteins with native disordered structure—named Intrinsically Disordered Proteins (IDPs)—were reported to be involved in numerous key biological processes, including cell cycle control, regulation, signaling and binding events involving biological macromolecules [1–5]. Intrinsic disorder depends on the type of amino acids composing a certain protein [6–8]. The structural plasticity of proteins is mainly induced by the presence of residues inherently flexible, polar and charged. Through inherent flexibility, the mentioned residues have the ability to recognize different ligands, including other proteins, membranes, and nucleic acids [8]. Further studies have also shown that a flexible intrinsically disordered region (IDR) in a given protein can fold in several ways as a consequence of its binding to distinct partners. Moreover, the residue flexibility may also allow different protein sequences to recognize the same binding site on a specific partner [9]. Often, IDPs or IDRs are critically involved in receptor signaling, where they modulate the native and functional state of many proteins [10]. However, there is a limit to the understanding of the molecular mechanisms that regulate biological functions in which IDPs are involved: there is a lack of knowledge about the physical basis on which these processes rely [11]. Recently, authors have demonstrated that each IDP dynamically explores an ensemble of unfolded configurations, then each IDP assumes more stable and ordered structures after binding to their ligands is complete [12,13]. Molecular dynamics (MD) simulations can provide useful information about these issues in the spatio-temporal resolution. In fact, the flexibility of a protein can be derived from the trajectory of a MD simulation by calculating the root mean-squared fluctuations (RMSFs) of single atoms after removing the translational and rotational movements. Vibrations around equilibrium are not random but depend on local structure flexibility. Hence, RMSFs capture the fluctuation of each atom around average positions. Usually, the analysis of the average atomic mobility of backbone atoms (N, C^{α} and C atoms) during MD simulation gives insight into the flexible and rigid regions of proteins [14].

In this paper we designed short peptides as IDRs model systems in order to better understand molecular recognition processes involving IDPs. In particular, we focused on the two following synthetic linear peptides: PepK: Y-G-E-C-P-C-K-OAllyl; PepE: Y-G-E-C-P-C-E-OAllyl.

Both these peptides were designed by choosing a few residues with disorder propensity. At the *C*-termini of PepK and PepE a Lys and Glu residues were added, respectively, aiming to study the effect of charges in modulating the binding properties on a general target. The reactive allylic ester group was added to easily allow further chemical modifications finalized to improve biological activities.

The conformational behavior of the two peptides was firstly analyzed by CD (Circular Dichroism) and 1D [^{1}H] and 2D [^{1}H, ^{1}H] NMR (Nuclear Magnetic Resonance) spectroscopies in order to establish their high flexibility and their inability to adopt a single ordered conformation. In addition, MD simulations indicated that the two peptides could be adequately represented by conformational ensembles.

In an attempt to define a biological significance for the above mentioned peptides, we have investigated their role as interactors of a known drug-target, *i.e.*, the CXCR4 receptor. Its overexpression results in metastatic dissemination of breast cancer cells to the lungs and lymph nodes and they contribute to melanoma tumor cell dissemination selectively to the lungs but not to the lymph nodes [15,16]. Moreover, CXCR4 has also been found to be a prognostic marker in multiple cancers including melanoma, colon cancer, leukemia, breast cancer, and gliomas [15,16]. Therefore, as CXCR4 is overexpressed in several human cancers, the blockade of CXCR4–CXCL12 interaction has been extensively investigated as a potential cancer therapeutic. Recently, we designed and synthesized two short flexible peptides, they both possess a CPC motif. Then we analyzed their structure by means of experimental and computational methods. Finally, we evaluated their interaction with CXCR4 by means of *in vitro* studies at 1 and 10 μM concentrations [15].

Hence, in this work we carried out biological tests to evaluate if PepE and PepK were able to: (i) interact with the CXCR4 receptor; (ii) inhibit CXCL12-induced migration; and (iii) reduce cAMP (Cyclic Adenosine Monophosphate) levels. Our data pointed out that PepE and PepK target the CXCR4 receptor by modulating the adenylate cyclase and they represent the starting point to design useful "*instruments*" for molecular recognition processes in which a flexible conformation is needed for the interaction with a specific target.

2. Results and Discussion

The two linear intrinsically disordered peptides PepK and PepE have been designed choosing a few amino acids with disorder propensity. A Lys residue has been added at *C*-terminal side of PepK and a Glu residue has been added at the *C*-terminal side of PepE. The conformational features of both peptides were studied in solution by CD, NMR and MD techniques.

2.1. Circular Dichroism

The secondary structure of PepE and PepK in PBS was studied by circular dichroism (CD) spectroscopy and analyzed by the CAPITO web server (http://capito.nmr.fli-leibniz.de). As expected for very short peptide sequences, both peptides did not show any tendency to fold under this experimental condition [17]. CD spectra (Figure 1) show a typical shape of an unordered structure with a negative band between 197 and 201 nm. This behavior is also confirmed by CAPITO analysis [18],

where spectra values—such as mean residue ellipticity (θ)—at λ = 200 nm are plotted *versus* λ = 222 nm. In fact, both the peptides are located in unfolded regions. It is important to notice that the CD spectrum of PepK presents also a maximum centered at 225 nm that has been already associated to the presence of fluctuating Polyproline II (PII) helices by various authors in recent papers [13,19,20].

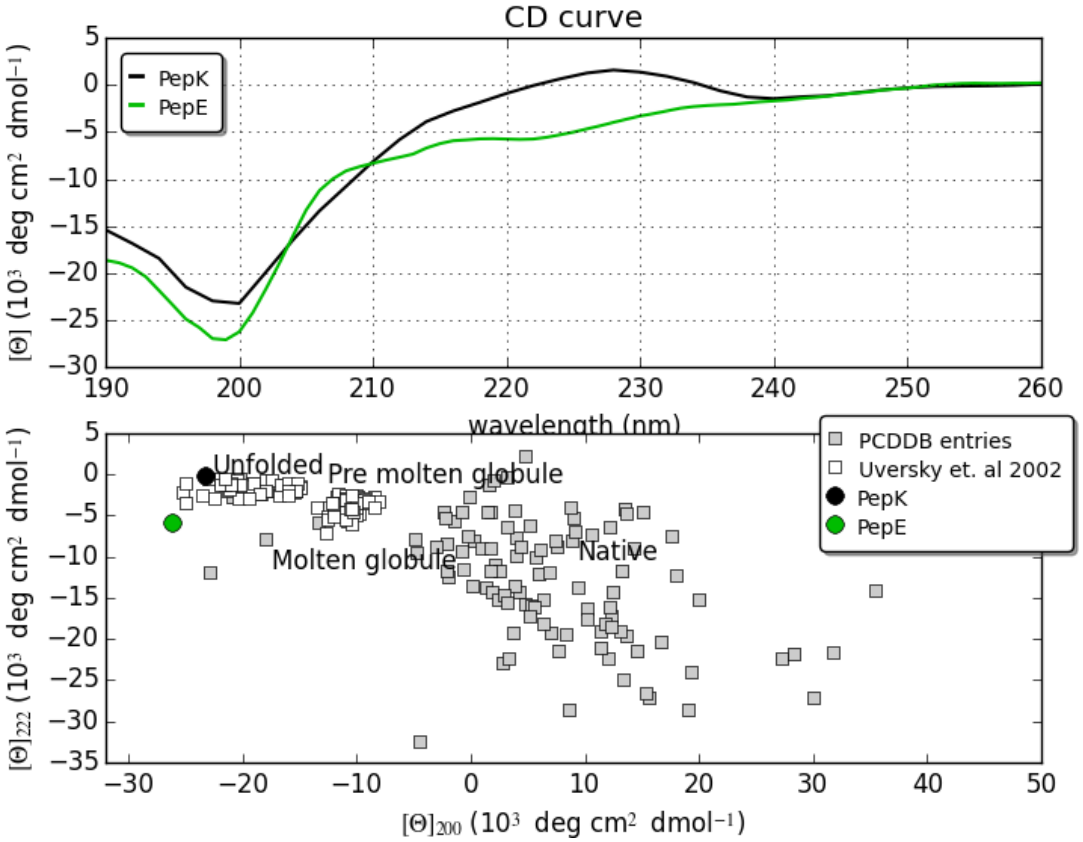


Figure 1. CD spectra of PepK and PepE (**upper** panel) and CD analysis by CAPITO tool (**lower** panel) where both peptides are indicated as unfolded peptides.

2.2. NMR Characterization

In aqueous solution, the poor spectral dispersion of the 1D [^{1}H] spectra and the almost complete absence of signal in the 2D [^{1}H, ^{1}H] NOESY [21] experiments indicated that both peptides were very flexible (data not shown). However, complete proton resonance assignments were achieved by combined analysis of the 2D [^{1}H, ^{1}H] TOCSY [22] and ROESY [23] spectra (Figures 2 and 3; Tables S1 and S2).

Moreover, we evaluated chemical shifts deviations of H_α protons from random coil values (Tables S3 and S4), which resulted in prevalence small and positive and characteristic of an extended disordered conformation [24]. The disorder state of both peptides was further strengthened by ROE patterns (Figure S1) which showed strong and sequential H_α^i–H_N^{i+1} contacts typical of flexible peptides [25]. Due to the proximity of chemical shifts between Cys4 H_α proton and water in both PepE and PepK, we could not clearly identify the configuration (*i.e.*, *cis* or *trans*) of the Cys4–Pro5 peptide bond.

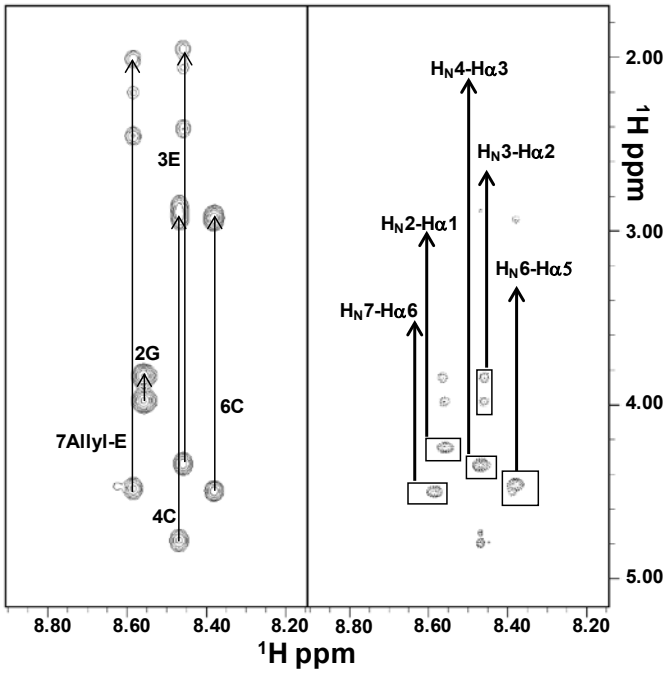


Figure 2. Comparison of 2D [^{1}H, ^{1}H] TOCSY (**left**) and ROESY (**right**) spectra of PepE in H_2O/D_2O (90/10). The H_N-aliphatic protons correlation regions are shown in each panel; spin system assignments are indicated in the left side. In the right panel, sequential ROE contacts are highlighted by rectangles and the corresponding assignments are indicated.

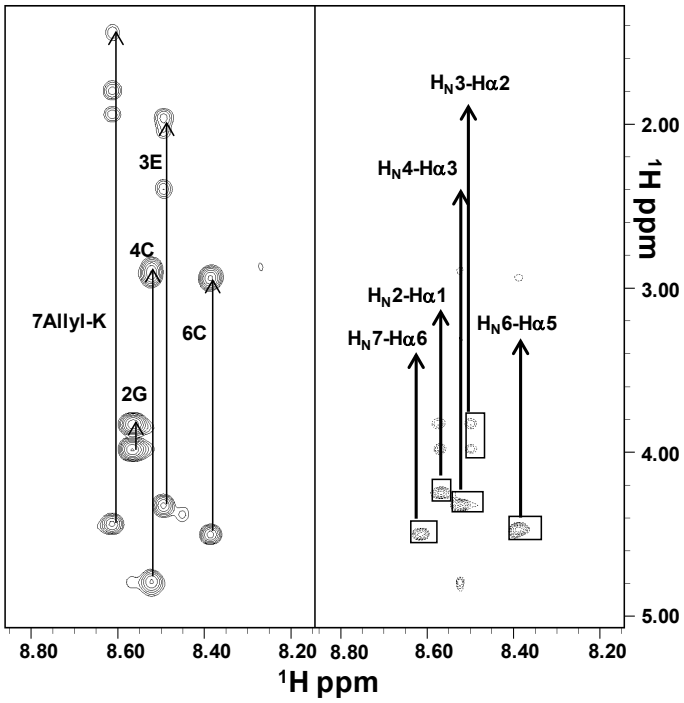


Figure 3. Comparison of 2D [^{1}H, ^{1}H] TOCSY (**left**) and ROESY 200 (**right**) spectra of PepK in H_2O/D_2O (90/10). The two sides of the figure show spectral regions containing H_N/aliphatic protons correlations; spin system assignments are indicated in the TOCSY left panel. In the right panel, sequential ROE contacts are shown.

2.3. Molecular Dynamics Simulations on the Two Peptides

Molecular dynamics simulations were performed on the two peptides, which were linearly modeled as reported in the Material and Methods Section, at neutral pH. In Figure 4A, we show the plots of the Root Mean Square Deviation (RMSD), computed by overlapping the various structures during simulations in respect to the initial conformation. RMSD plots of PepK and PepE evidence high levels of fluctuation, thus suggesting that these peptides are flexible. On the one hand, this result is confirmed also by Root Mean Square Fluctuation (RMSF) plots where the residues located at *N*- and *C*-termini present higher values of RMSF (Figure 4B). Instead, the gyration radii decreased during the simulations (Figure 4C) reaching a value of 0.5 nm. This suggests an increasing compactness of the two peptides, which result as if they were stabilized by a certain number of H-bonds of main chain–main chain (MM), main chain–side chain (MS) and side chain–side chain (SS) types (Figure 4D). For this reason, the two peptides tended to become more compact and the radius of gyration decreased. According to the data obtained by CD and NMR measurements in aqueous solution, the analysis of the secondary structure evolution clearly showed the high flexibility of both peptides during the simulations, which makes difficult the formation of stable regular secondary structure elements, *i.e.*, helix and β-strand. Very flexible peptide structure are normally characterized by several conformers at equilibrium, hence they could be well described by a conformational ensemble. Having taken this property into proper account, we have firstly performed a cluster analysis to determine the groups of structures that share similar conformational features according to their RMSD values.

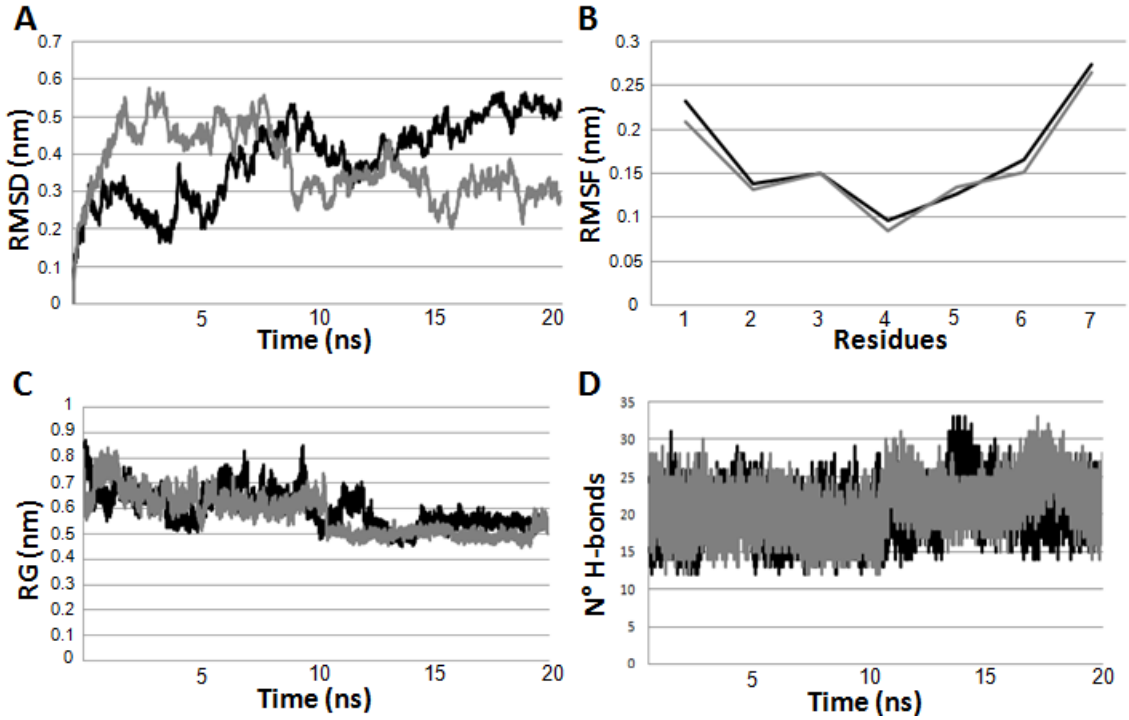


Figure 4. Analysis of the molecular dynamics simulations performed on PepE (in black) and PepK (in grey) at physiological pH in explicit solvent in terms of: (**A**) root mean square deviation (RMSD) plot; (**B**) root mean square fluctuation (RMSF); (**C**) gyration radius plot; and (**D**) H-bonds plot.

The number of the most populated clusters resulted as equal to 8 for PepE and 9 for PepK. These clusters are mainly stabilized by MM H-bonds (Figure 5) that involved the atomic groups of the peptide backbone. This structural organization was observed by Pappu *et al.* [26] for charged peptides: it closely resembles the organization of collapsed and slightly soluble globules.

Since CD spectra of PepK pointed out the presence of PII, we verified its possible occurrence in the most populated clusters by using the angular ranges following reference parameters for the angles: $-110° \leq \varphi \leq -40°$ and $130° \leq \psi \leq 180°$ in the Ramachandran map [27]. No residues in PII were present in eight clusters of PepE whereas in 5 out of 9 clusters of PepK we identified residues in PII (Figure S2), according to CD analysis.

It is well known that the peptide dynamics depend on the surrounding solvent, which mediates interactions among residues. Therefore, we have analyzed the total number of water molecules in the system and their possible role [28]: the simulation box for PepE and PepK contained 2956 and 2611 water molecules, respectively. The total average number of water molecules that formed H-bonds resulted slightly higher for PepE if compared to PepK: 33 and 30 in PepE and PepK, respectively. This trend could be due to the different amino acid sequence of the two peptides; in fact PepK has got two oppositely charged residues in 3 and 7 positions, hence it may assume a more compact structure. The majority of H-bonds (indicatively 60%) for both peptides resulted to involve the peptide backbone, being of main chain—water oxygen type (MH) or of the water oxygen and main chain type (HM). A detailed analysis of the residues involved in H-bonds with water molecules pointed out that the residues in positions 1, 3 and 7 (Y, E and E in PepE and Y, E and K in PepK) have developed a higher number of H-bonds with water molecules than the residues located in the other four positions (Table S5). Furthermore, the average number of H-bonds developed by E and K with water molecules was lower than the average number of the H-bonds developed by E and E with water molecules in PepE. Hence, these results confirmed that PepE tended to assume more extended conformations being that it contains two negatively charged residues that could allow large interactions with water molecules.

Generally, the data evidence that the structure of both peptides is very flexible, and can be dynamically stabilized by a network of MM H-bonds. Hence, their structures can be adequately represented as conformational ensembles characterized by fluctuating irregular secondary structures. In fact, the peptides do not assume stable and fixed conformations: they go continuously from one cluster to another. As the high flexibility of peptides did not allow us to calculate a 3D model based on NMR parameters, the MD was very useful to understand how these peptides can move as well as to understand that the best conformers needed to be put together to have a clear representation. In particular, the two peptides do not explore the full space, but rather, the rapid interchange among the flexible conformers induce quick changes between clusters. This is in line with the data shown in Figure 4B where it is visible that the more flexible residues are located in the terminal ends.

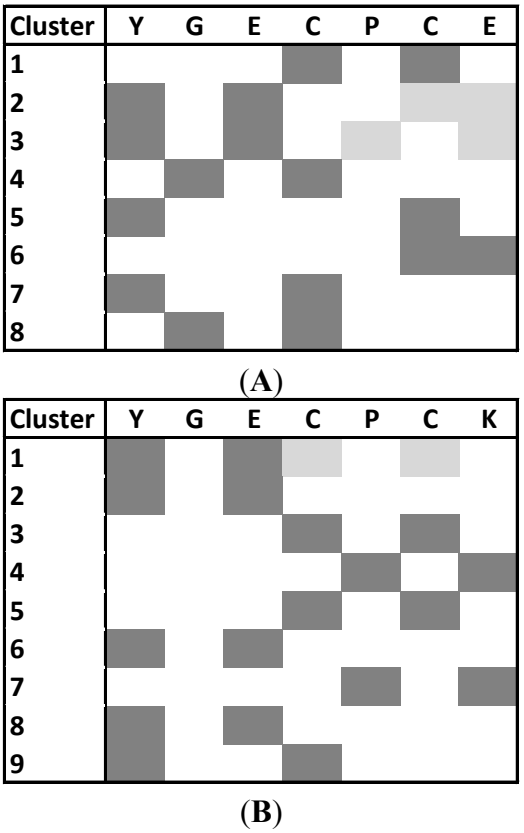


Figure 5. The map of MM H-bonds in the 8 and 9 clusters calculated for the PepE (**A**) and PepK (**B**), respectively, during MD simulations. We reported with the same colors the residues involved in the same MM H-bonds.

2.4. Functional Characterization on CXCR4 Receptor

We have designed a short disordered peptide the containing CPC motif and we have analyzed its binding capability on the CXCR4 receptor finding a behavior as ligand [15]. Hence, we performed PepK and PepE functional characterization on the same chemokine receptor. In order to assess the inhibitory capability of PepK and PepE on CXCR4 function, we have evaluated peptides binding to the receptor through an indirect binding assay; then, in addition we have assessed the inhibition of CXCL12-induced migration and evaluated the activity of the adenylate cyclase function. The results were compared to Plerixafor (known as AMD3100) that is used as a clinical antagonist of CXCR4. The treatment with plerixafor plus Granulocyte-colony stimulating factor (G-CSF) induced the mobilization of hematopoietic precursors from bone marrow, allowing a more efficient recovery of the CD34 stem cells necessary for autologous stem cell transplantation [29].

The binding of the two peptides to CXCR4 was evaluated in CCRF-CEM cells using a CXCR4 12G5 monoclonal antibody as previously described [30]. PepK and PepE were evaluated at concentrations similar to the peptidic antagonists recently developed [16,30]. Indeed, we have tested only two peptide concentrations (1 and 10 μM), which are matching the ones previously used in CXCR4 binding assays. As shown in Figure S3, PepK and PepE did not displace the 12G5 antibody compared to AMD3100, which strongly reduced antibody binding to CXCR4 (Figure S3). These results may be explained considering the different binding sites on AMD3100/12G5 antibody and the high flexibility of the two

disordered peptides. Moreover, cAMP and migration assays were carried out as well in order to evaluate the peptide ability to inhibit the activation of CXCR4 mediated intracellular pathways.

CXCR4 is a G-Protein coupled receptor, which, as result of the binding to CXCL12, inhibits through the G-protein the adenylate cyclase activity and the inhibition of cAMP production as "second messenger" [31]. In Figure S4, preliminary biological results are summarized. The results underline that in the presence of CXCL12 (100 ng/mL) and Forskolin (1 μM): (i) PepE determines a dose dependent increase of the cAMP level of 20% at the lower dose (1 μM) and 65% at the higher dose (10 μM); (ii) PepK increases cAMP but the effect is not dose-dependent; and (iii) AMD3100 has activities comparable or even less efficient than PepE. Hence, PepK and PepE result to modulate the adenylate cyclase at equal or higher levels as AMD3100.

Finally, the migration was performed (Figure S5) and, both PepK and PepE reduced CCRF-CEM cells migration towards CXCL12 although less effectively than AMD3100. In particular, PepE reduces the migration index from 1.75- to 1.35-fold and PepK from 1.75- to 1.55-fold, whereas AMD3100 from 1.75- to 1.15-fold.

3. Experimental Section

Protected fluorenylmethoxycarbonyl (Fmoc)N^{α}-Fmoc-amino acid derivatives, 2-chlorotrityl chloride resin pre-loaded with Fmoc-Glu-OAllyl and Fmoc-Lys-OAllyl and coupling reagents were purchased from Novabiochem (Schwalbach, Germany) and Iris Biotech GMBH (Marktredwitz, Germany). All other chemicals were commercially available by Sigma–Aldrich (Milwaukee, WI, USA) and Acros (Geel, Belgium). The purity of products was determined by analytic reverse-phase HPLC using a VWR Hitachi instrument (Radnor, PA, USA) equipped with a L-2450 auto sampler, two L-2130 pumps, a Satisfaction RP18-AE column (5 μm, 250 × 4.6 mm) and a L-2450 diode array detector, at a flow rate of 0.8 mL/min.

The compounds were purified by preparative reverse-phase HPLC using a VWR LaPrep system consisted of a P202 injector, two P110 pumps, a Satisfaction RP18-AB C18 column (5 μm, 250 × 20 mm) and a P314 UV detector, at a flow rate of 10 mL/min.

The following eluents were used in a gradient mode: (A) 0.1% trifluoroacetic acid (TFA) in H_2O/CH_3CN (95/5) and (B) 0.1% TFA in CH_3CN/H_2O (95/5). Water was of Milli-Q quality and was obtained after filtration of distilled water through a Milli-Q® cartridge system (Millipore Co., Billerica, MA, USA). CH_3CN and TFA were of HPLC use quality. Degassing of solvents was performed using argon bubbling. Gradient used for the analytic and preparative RP-HPLC was respectively from 0% B to 100% B in 15 min and from 0% to 70% B in 70 min; UV detection at 214 nm. CH_3CN was evaporated and the aqueous solution was freeze-dried to give purified PepK and PepE as white solids (68% and 54% yield, respectively).

Peptide identity was confirmed by mass spectrometry analyses performed on an Ultraflex III TOF/TOF system (Bruker Daltonics, Bremen, Germany), equipped with 200 Hz smart beam laser (355 nm) and operating in reflectron positive ion mode. Mass spectra were acquired over the *m*/*z* range 450–5000 by accumulating data from 1000 laser shots for each spectrum. The instrumental conditions employed to analyze molecular species were the following: ion source 1: 25.08 kV; ion source 2: 21.98 kV, lens: 11.03 kV, pulsed ion extraction: 30 ns, reflector: 26.39 kV, reflector 2: 13.79 kV. Matrix

suppression was activated by deflection mode: suppression up to 450 Da. Mass calibration was performed for each samples in range of ~700–3200 Da with a peptide calibration mixture (8206195, Peptide Calibration Standard, Bruker Daltonics). The instrument was controlled using Bruker's flexControl 3.4 software and mass spectra were analyzed in Bruker's FlexAnalysis 3.4 software.

3.1. Synthesis

PepK (sequence: YGECPCK-OAllyl) and PepE (sequence: YGECPCK-OAllyl) were synthesized by Fmoc standard chemistry protocol on 2-chlorotrityl chloride resin pre-loaded with the side chain of a Fmoc-Lys-OAllyl and a Fmoc-Glu-OAllyl, respectively (0.24 mmol scale).

Swelling and Fmoc-deprotection: Swelling and Fmoc-deprotection steps were performed by treating the resin for 30 min in dichlorometane (DCM) and in dimethylformamide (DMF)/Piperidine (Pip) (80/20, *v*/*v*) solution, respectively.

Washing: Washing steps after coupling and deprotection steps were performed by treating resin in DMF (3 × 1 min), MeOH (1 × 1 min), and DCM (3 × 1 min), successively.

Amino acid coupling: The resin was immersed and mixed in a DMF solution containing the Fmoc-protected amino acid (3 eq), *O*-benzotriazole-tetramethyl-uronium-hexafluoro-phosphate (HBTU) (3 eq), and diisopropylethylamine (DIEA) (6 eq) for 2 h at room temperature.

Peptide cleavage: The peptide was cleaved by immersing the resin in TFA/H_2O/triisopropylsilane (TIS) (30 mL, 95/2.5/2.5, *v*/*v*/*v*) for 3 h. The cleavage cocktail was filtered from the resin, peptides were precipitated with ice-cold diethyl ether, centrifuged, and decanted. Precipitation, centrifugation, and decantation operations were repeated twice. The resulting white solid was dissolved in water (10 mL) and freeze-dried to give a white powder that was analyzed for purity and purified by preparative RP-HPLC, using the specified conditions.

3.2. Circular Dichroism Measurement

Far-UV CD spectra were recorded from 190 to 260 nm on a Jasco J-810 spectropolarimeter (JASCO Inc., Easton, MD, USA) equipped with a NesLab RTE111 thermal controller unit using a 1 mm quartz cell at 25 °C. Circular dichroism measurements were carried out on lyophilized peptides dissolved in 10 mM phosphate buffer at pH 7.4 M at a peptide concentration of 5×10^{-6} M. Other experimental settings were: scan speed, 10 $nm \cdot min^{-1}$; sensitivity, 50 mdeg; time constant, 16 s; bandwidth, 1 nm. Each spectrum was obtained averaging three scans, and by subtracting contributions from other species in solution and converting the signal to mean residue ellipticity in units of $deg \cdot cm^2 \cdot dmol^{-1} \cdot res^{-1}$.

3.3. NMR Analysis

NMR spectra were recorded at 25 °C on a Varian Unity Inova 600 MHz spectrometer (Varian Inc., Palo Alto, CA, USA) provided with a cold probe. The process of proton resonance assignments was carried out with a canonical protocol [25] based on analysis of the following two dimensional [1H, 1H] spectra: TOCSY (Total Correlation Spectroscopy) (70 ms mixing time) [22], DQFCOSY (Double Quantum Filter Correlation Spectroscopy) [32], NOESY (Nuclear Overhauser Enhancement

Spectroscopy) [21] (300 ms mixing time) and ROESY (Rotating frame Overhauser Enhancement Spectroscopy) [23] (200 ms mixing time). Chemical shifts were referenced with respect to the TSP (trimethylsilyl-3-propionic acid sodium salt-d4, 99% d, (Armar Scientific, Döttingen, Switzerland) signal at 0.0 ppm.

1D spectra were acquired with a relaxation delay of 1 s and 32–128 scans. 2D experiments were generally acquired with 32–64 scans, 128–256 FIDs in t_1, 1024 or 2048 data points in t_2. The DPFGSE (Double Pulsed Field Gradient Selective Echo) sequence [33] was used to suppress water signal. Spectra were processed with the Varian software VNMRJ 1.1D (Varian by Agilent Technologies, Milan, Italy) and analyzed with the NEASY [34] program that is included in the CARA (Computer Aided Resonance Assignment) software package (http://www.nmr.ch/).

3.4. Molecular Modeling and Dynamics Simulations

PepK and PepE have been built by using the Builder module in InsightII and then subjected to molecular dynamics (MD) simulations performed with the GROMACS software package (v3.3.1) [35]. In detail, each peptide was put in a cubic box filled with SPC216 water molecules, and GROMOS43a1 was selected as force field because it is commonly used as force field for MD simulations on peptides and our group has already used it in other previous papers in which we reported some conformational studies on different peptides where the results obtained from MD simulations proved to be in excellent agreement with CD and/or NMR studies [13,15,16]. In order to optimize the systems, the peptides were previously subjected to energy minimization and position restraints cycles. The simulations were carried out with periodic boundary conditions by adding chloride ions so that the net electrostatic charge of the system was zero. The bond lengths were constrained by the linear constraint solver algorithm. Particle mesh Ewald algorithm was used for the electrostatic interactions with a cutoff of 0.9 nm, according to our recent papers [13,15,36]. All simulations were run for 20 ns at neutral pH and room temperature (300 K) coupling to the system an external bath. GROMACS routines RMSD and RMSF, gyration radius, number of H-bonds and secondary structure evolution were utilized to check the trajectories and the quality of the simulations. Additionally, on the basis of the distances between structures, like RMSD values, we founded a set of clusters that reflects the range of conformations accessible and the relative weight of each of these by using a clustering algorithm implemented in GROMACS as reported also in our recent papers [13,16]. The presence of putative H-bonds between the residues and with water molecules was evaluated by Hbplus [37].

3.5. Migration Assay

CCRF-CEM cells migration was assayed in 24-well Transwell chambers (Corning Inc., Corning, NY, USA) using inserts with an 8-μm pore membrane. Membranes were precoated with collagen (human collagen type I/III) and fibronectin (20 mg/mL each). CCRF-CEM cells were placed in the upper chamber (1×10^5 cells/well) in RPMI containing 1% BSA (migration media). Cells were pre-incubated for 45 min with CXCR4 antagonist and allowed to migrate toward 100 ng/mL CXCL12 in the lower chamber. After 16 h incubation, migrated cells were collected from the lower chamber and counted. The migration index was defined as the ratio between migrating cells in the experimental group and migrated cells in the control group.

3.6. Binding Assay

PepK and PepE binding to CXCR4 was evaluatedas previously described [30]. Briefly 2.5×10^5 CCRF–CEM cells were pre-incubated with 10 μM antagonist peptides in binding buffer (PBS 1× plus 0.2% BSA and 0.1% NaN_3) for 1 h at 37 °C, 5% CO_2 and then labeled for 45 min with anti-CXCR4 PE-antibody (FAB170P, clone 12G5, R&D Systems, Minneapolis, MN, USA). Cells were washed in PBS and analyzed by FACS Canto II cytofluorimeter (Becton Dickinson Immunocytometry Systems, Mountain View, CA, USA).

3.7. cAMP Assay

CCRF-CEM cells (1×10^6) were incubated in presence of PepK, PepE or Plerixafor (known as AMD3100) being a CXCR4 antagonist that has provided proof of concept for inhibition of the pathway [31] at different concentrations (1 and 10 μM) in combination with forskolin (F) (1 μM) for 20 min, followed by stimulation with CXCL12 (100 ng/mL) for 10 min. Controls include cells stimulated with CXCL12 and forskolin or forskolin alone in absence of anti-CXCR4 inhibitors. Cells are harvested and lysed with 0.1 M HCl and cAMP levels was assayed using a direct competitive enzyme immunoassay (BioVision Incorporated, Milpitas, CA, USA) according to manufacture instructions.

4. Conclusions

As CXCR4 is overexpressed in several human cancers, the blockade of CXCR4–CXCL12 interaction represents from many years an interesting drug-target [15]. Moreover, the plerixafor (previously known as AMD3100) is the most clinically advanced compound even if it showed cardiotoxicity, as reported in its clinical trial against HIV [38]. Therefore, this suggests the need to develop new antagonists able to block, without any other compliance, the binding between CXCL12 and CXCR4. In this work we have described the structural preferences in solution of two synthetic linear peptides containing a few amino acids with disorder propensity, and an allylic ester group at the *C*-terminal end. NMR and CD solution studies, revealed the absence of regular secondary structure elements thus confirming the natively disordered nature of these peptides. MD simulations evidenced that, even if they are flexible, they are stabilized by a network of transient and dynamic intra-molecular H-bonds of MM type and by interactions with water molecules. In details, NMR and computational studies showed the flexibility of PepE and PepK, their inability to adopt a single ordered conformation and, hence, the need to represent them by conformational ensembles characterized by the presence of transient and dynamic MM H-bonds.

Moreover, we have performed a functional characterization of the two peptides by investigating (a) their ability to interact with the CXCR4 receptor; (b) the inhibition of CXCL12-induced migration and (c) the reduction of cAMP levels. Generally, we pointed out that PepK and PepE result to modulate the adenylate cyclase as much or more as AMD3100, *i.e.*, the best characterized CXCR4 inhibitor. These results indicate that the short flexible peptides could efficiently target the CXCR4 receptor.

This preliminary evidence suggests that the PepK and PepE interact with CXCR4 and they represent a potential starting system able to be converted to more efficient *in vitro* and *in vivo* antagonists of CXCR4.

Supplementary Materials

Supplementary materials are available at http://www.mdpi.com/1422-0067/16/06/12159/s1.

Acknowledgments

Marian Vincenzi thanks the Università Italo Francese (UIF) for financial support, project C2-1, 2014/38723, cap. 6.01.1810 UIF. Luisa Ronga and Jean Guillon thank the "Plateforme Protéome" of the University of Bordeaux for the access to the Ultraflex III TOF/TOF mass spectrometer and the precious advice in using it. The authors thank Leopoldo Zona and Luca De Luca for technical assistance.

Author Contributions

Marian Vincenzi and Marilisa Leone carried out NMR experiments and participated in writing the manuscript. Luisa Ronga and Jean Guillon performed peptide synthesis, purification and chemical physical characterization of linear peptides. Diego Tesauro and Antonella Accardo recorded and analyzed CD spectra. Susan Costantini and Giovanni Colonna performed MD analysis and contributed to write the paper. Stefania Scala, Luigi Portella and Anna Maria Trotta performed the biological test. Filomena Rossi and Luisa Ronga designed the research and wrote the paper. All authors have read and approved the final manuscript.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Habchi, J.; Tompa, P.; Longhi, S.; Uversky, V.N. Introducing protein intrinsic disorder. *Chem. Rev.* **2014**, *114*, 6561–6588.
2. He, B.; Wang, K.; Liu, Y.; Xue, B.; Uversky, V.N.; Dunker, A.K. Predicting intrinsic disorder in proteins: An overview. *Cell Res*. **2009**, *19*, 929–949.
3. Motlagh, H.N.; Wrabl, J.O.; Li, J.; Hilsen, V.J. The ensemble nature of allostery. *Nature* **2014**, *508*, 331–338.
4. Wright, P.E.; Dyson, H.J. Intrinsically unstructured proteins: Re-assessing the protein structure-function paradigm. *J. Mol. Biol.* **1999**, *293*, 321–331.
5. Iakoucheva, L.M.; Brown, C.J.; Lawson, J.D.; Obradovic, Z.; Dunker, A.K. Intrinsicdisorder in cell-signaling and cancer-associated proteins. *J. Mol. Biol.* **2002**, *323*, 573–584.
6. Schweitzer-Stenner, R. Conformational propensities and residual structures in unfolded peptides and proteins. *Mol. BioSyst.* **2012**, *8*, 122–133.
7. Uversky, V.N. Intrinsically disordered protein from A to Z. *Int. J. Biochem. Cell Biol.* **2011**, *43*, 1090–1103.
8. Leone, M.; Mercurio, F.A.; Vincenzi, M.; Accardo, A.; Ringhieri, P.; Tesauro, D.; Carrière, F.; Rossi, F. Conformational disorder in phosphopeptides: Solution studies by CD and NMR techniques. *Peptidomics* **2014**, *1*, 14–21.

9. Miller, M. The importance of being flexible: The case of basic region leucine zipper transcriptional regulators. *Curr. Protein Pept. Sci.* **2009**, *10*, 244–269.
10. Sigalov, A.B.; Uversky, V.N. Differential occurrence of protein intrinsic disorder in the cytoplasmic signaling domains of cell receptors. *Self Nonself* **2011**, *2*, 55–72.
11. Costantini, S.; Sharma, A.; Raucci, R.; Costantini, M.; Autiero, I.; Colonna, G. Genealogy of an ancient protein family: The Sirtuins, a family of disordered members. *BMC Evol. Biol.* **2013**, *13*, 60–80.
12. Raucci, R.; Colonna, G.; Giovane, A.; Castello, G.; Costantini, S. *N*-terminal region of human chemokine receptor CXCR3: Structural analysis of CXCR3(1–48) by experimental and computational studies. *Biochim. Biophys. Acta* **2014**, *1844*, 1868–1880.
13. Mittal, J.; HyeonYoo, T.; Georgiou, G.; Truskett, T.T. Structural ensemble of an intrinsically disordered polypeptide. *J. Phys. Chem. B* **2013**, *117*, 118–124.
14. Chen, S.; Gao, S.; Cheng, D.; Huang, J. The characterization and comparison of amyloidogenic segments and non-amyloidogenic segments shed light on amyloid formation. *Biochem. Biophys. Res. Commun.* **2014**, *447*, 255–262.
15. Costantini, S.; Raucci. R.; Colonna, G.; Mercurio, F.A.; Trotta, A.M.; Ringhieri, P.; Leone, M.; Rossi, F.; Pellegrino, C.; Castello, G.; *et al.* Peptides targeting chemokine receptor CXCR4: Structural behavior and biological binding studies. *J. Pept. Sci.* **2014**, *20*, 270–278.
16. Palladino, P.; Portella, L.; Colonna, G.; Raucci, R.; Saviano, G.; Rossi, F.; Napolitano, M.; Scala, S.; Castello, G.; Costantini, S. The *N*-terminal region of CXCL11 as structural template for CXCR3 molecular recognition: Synthesis, conformational analysis, and binding studies. *Chem. Biol. Drug Des.* **2012**, *80*, 254–265.
17. Underfriend, S.; Meienhofer, J. *In The Peptides*; Hruby, V.J., Ed.; Academic Press: New York, NY, USA, 1985; Volume 7.
18. Wiedemann, C.I.; Bellstedt, P.; Görlach, M. CAPITO-a web server-based analysis and plotting tool for circular dichroism data. *Bioinformatics* **2013**, *29*, 1750–1757.
19. Sreerama, N.; Woody, R.W. Poly(Pro)II helices in globular proteins: Identification and circular dichroic analysis. *Biochemistry* **1994**, *33*, 10022–10025.
20. Bhatnagar, R.S.; Gough, C.A. *Circular Dichroism and the Conformational Analysis of Biomolecules*; Plenum Press: New York, NY, USA, 1996.
21. Kumar, A.; Ernst, R.R.; Wuthrich, K. A two-dimensional nuclear Overhauser enhancement (2D NOE) experiment for the elucidation of complete proton-proton cross-relaxation networks in biological macromolecules. *Biochem. Biophys. Res. Commun.* **1980**, *95*, 1–6.
22. Griesinger, C.; Otting, G.; Wüthrich, K.; Ernst, R.R. Clean TOCSY for proton spin system identification in macromolecules. *J. Am. Chem. Soc.* **1988**, *110*, 7870–7872.
23. Bax, A.; Davis, D.G. Practical aspects of two-dimensional transverse NOE spectroscopy. *J. Magn. Reson*. **1985**, *63*, 207–213.
24. Wishart, D.S.; Sykes, B.D.; Richards, F.M. Relationship between nuclear magnetic resonance chemical shift and protein secondary structure. *J. Mol. Biol.* **1991**, *222*, 311–333.
25. Wüthrich, K. *NMR of Proteins and Nucleic Acids*; John Wiley & Sons: New York, NY, USA, 1986.

26. Das, R.K.; Pappu, R.V. Conformations of intrinsically disordered proteins areinfluenced by linear sequence distributions of oppositely charged residues. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 13392–13397.
27. Vila, J.A.; Baldoni, H.A.; Ripoll, D.R.; Ghosh, A.; Scheraga, H.A. PII helix conformation in a proline-rich environment: A theoretical study. *Biophys. J.* **2004**, *86*, 731–742.
28. Tarek, M.; Tobias, D.J. Role of protein–water hydrogen bond dynamics in the protein dynamical transition. *Phys. Rev. Lett.* **2002**, *88*, 138101.
29. De Clercq, E. AMD3100 story: The path to the discovery of a stem cell mobilizer (Mozobil). *Biochem. Pharmacol.* **2009**, *77*, 1655–1664.
30. Portella, L.; Vitale, R.; de Luca, S.; D'Alterio, C.; Ieranò, C.; Napolitano, M.; Riccio, A.; Polimeno, M.N.; Monfregola, L.; Barbieri, A.; *et al.* Preclinical development of a novel class of CXCR4 antagonist impairing solid tumors growth and metastases. *PLoS ONE* **2013**, *8*, e74548.
31. Teicher, B.A.; Fricker, S.P. CXCL12 (SDF-1)/CXCR4 pathway in cancer. *Clin. Cancer Res.* **2010**, *16*, 2927–2931.
32. Piantini, U.; Sorensen, O.W.; Ernst, R.R. Multiple quantum filters for elucidating NMR coupling networks. *J. Am. Chem. Soc.* **1982**, *104*, 6800–6801.
33. Dalvit, C. Efficient multiple-solvent suppression for the study of the interactions of organic solvents with biomolecules. *J. Biomol. NMR* **1998**, *11*, 437–444.
34. Bartels, C.; Xia, T.H.; Billeter, M.; Guntert, P.; Wuthrich, K. The program XEASY for computer-supported NMR spectral analysis of biological macromolecules. *J. Biomol. NMR* **1995**, *6*, 1–10.
35. Van Der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A.E.; Berendsen, H.J. GROMACS: Fast, flexible, and free. *J. Comput. Chem.* **2005**, *26*, 1701–1718.
36. Guariniello, S.; Colonna, G.; Raucci, R.; Costantini, M.; di Bernardo, G.; Bergantino, F.; Castello, G.; Costantini, S. Structure-function relationship and evolutionary history of the human selenoprotein M (SelM) found over-expressed in hepatocellular carcinoma. *Biochim. Biophys. Acta* **2014**, *1844*, 447–456.
37. McDonald, I.K.; Thornton, J.M. Satisfying hydrogen bonding potential in proteins. *J. Mol. Biol.* **1994**, *238*, 777–793.
38. Crump, M.P.; Gong, J.H.; Loetscher, P.; Rajarathnam, K.; Amara, A.; Arenzana-Seisdedos, F.; Virelizier, J.L.; Baggiolini, M.; Sykes, B.D.; Clark-Lewis, I. Solution structure and basis for functional activity of stromal cell derived factor-1; dissociation of CXCR4 activation from binding and inhibition of HIV-1. *EMBO J.* **1997**, *16*, 6996–7007.

Intrinsically disordered amphiphilic peptides: a new class of potential targets in drug discovery.

Marian Vincenzi [a,b,e#], Antonella Accardo[a,b#], Susan Costantini[c], Stefania Scala[d], Luigi Portella[d], Annamaria Trotta[d], Luisa Ronga[e], Jean Guillon[e], Marilisa Leone[b], Giovanni Colonna[c], Filomena Rossi[a], and Diego Tesauro[a,b*]

[a]*Department of Pharmacy and CIRPeB University of Naples "Federico II", Via Mezzocannone 16, I-80134 Naples, Italy;*

[b]*Istituto di Biostrutture e Bioimmagini – CNR, Via Mezzocannone 16, 80134, Naples, Italy Via Mezzocannone 16, 80134, Naples, Italy*

[c]*CROM, Istituto Nazionale Tumori "Fondazione G. Pascale", IRCCS, 80131 Napoli, Italy*;

[d]*Molecular Imunology and Immuneregulation, Istituto Nazionale Tumori "Fondazione G. Pascale", IRCCS , I-80131Napoli, Italy;*

[e] *UFR des Sciences Pharmaceutiques - Collège Sciences de la Santé, INSERM U869, Laboratoire ARNA,Université de Bordeaux, 146 rue Léo Saignat, 33076 Bordeaux cedex –France*

[fe]*Servizio di Informatica Medica, Azienda Ospedaliera Universitaria, Seconda Università di Napoli, I-80138 Napoli, Italy*;

s.costantini@istitutotumori.na.it, *s.scala@istitutotumori.na.it*, *portella@gmail.com*, *amt78@libero.it*, giovanni.colonna@unina2.it, *luisa.ronga@u-bordeaux.fr*, *jean.guillon@u-bordeaux.fr*

# These Authors contributed equally to this work

[*] Corresponding author. Tel.: +390812536643; fax: +390812534574.
E-mail : diego.tesauro@unina.it

ABSTRACT

In this paper we report the conformational analysis in solution by CD, fluorescence and DLS spectroscopies of functionalized linear peptides corresponding to sequences: Y-G-E-C-P-C-K-OAllyl and Y-G-E-C-P-C-E-OAllyl, henceforth called PepE and PepK respectively, designed by selecting amino acids with disorder propensity. The formation of stable supramolecular aggregates in aqueous solution can be promoted by conjugating two aliphatic chains of eighteen carbon atoms at the N-terminus of both sequences achieving two PAs ($(C18)_2$PepK and $(C18)_2$PepE). Two others PAs were got inserting a unit of 21-amino-4,7,10,13,16,19-hexaoxaheneicosanoic acid(AhOh) between the peptide sequence and the hydrophobic double-tail. This moiety plays a double task like spacer and linker. First of all the ethoxylic groups increase the hydrophilicity without modifying the final charge of the molecule thus avoiding intra and inter monomers electrostatic interactions then, divide the bioactive peptide from the hydrophobic shell. The relative short length of this moiety was determinate in order to avoid possible formation of a hydrophilic pocket that could hide the amino acid sequences. Biological properties related to recognition in cell lines overexpressing the CXCR4 receptor, were analyzed and discussed in comparison with previously studied disordered linear peptides.

Keywords:

Intrinsically disordered peptides (IDP) , Peptide Amphiphiles (PAs), Supramolecular aggregates, Circular Dichroism (CD), CXCR4 receptors, DLS

1. Introduction

Peptides may carry out a large array of biological functions acting as growth factors, neurotransmitters, and hormones.[1-5] The regulatory role played in physiological tasks, push pharmacologist to exploit peptides as therapeutic [6-7] or diagnostic tools.[8-9] Peptide drugs can offer many advantages: they are specific but have a broad spectrum of activity, they own less side effects than organic molecules; furthermore infinite sequence possibilities can potentially be designed to regulate physiological processes.[10] Moreover peptides are able to act as excellent building blocks for a wide range of biomaterials as they can be combined with a range of molecules such as lipids, sugars, nucleic acids, metallic nanocrystals and so on [11 -12] However one hurdle in development of peptide drugs is that they can elicit an immune response against themselves and furthermore are vulnerable to proteolytic enzymes. Many strategies could be used to overcome this drawbacks by modifying peptide sequences. One of them can be to bind one terminal moiety to surfaces of nanostructures or of supramolecular aggregates.[13] Among potential targets, peptides can act as agonist or antagonist by interacting with cell surface receptor systems that have been shown to be overexpressed in a variety of neoplastic and non-neoplastic cells.[14] Molecular Dynamics (MD) techniques can drive Peptide rational design efforts and help obtaining molecules in the most favorable binding conformation by introducing appropriate constraints.

The ability of a peptide to fold or not to fold under physiologic conditions is encoded in its amino acid sequence.[15-16] Recent literature data [17-18] have reported that each intrinsically disordered protein (IDP) dynamically searches an ensemble of unfolded configurations, assuming more stable ordered structures only after binding to their ligands. The ability of peptides to self-assemble is helped by the structural and chemical compatibilities among each molecule. A great improvement of using self-assembling peptides to build nanostructures is that specific features can be incorporated thus allowing peptide modifications and different functionalizations[12]. Recently we have analyzed the structural preferences in solution of two synthetic linear ester peptides encompassing the following

sequences: Y-G-E-C-P-C-K-OAllyl (PepK) and Y-G-E-C-P-C-E- OAllyl (PepE) characterized by the presence of a few amino acids able to promote disorder.[19] In our studies, NMR and CD spectra have revealed the lack of well-defined secondary structure elements thus confirming the conformational flexibility of both peptides . Concerning that PepK and PepE contained a CPC motif as another short disordered peptide that resulted to have a binding capability on CXCR4 receptor [20], we tested the biological properties of these peptides by an indirect binding assay on CXCR4 and by verifying the inhibition of CXCL12-induced migration and the reduction of cAMP levels. In details, CXCR4 is a rhodopsin-like seven-transmembrane G-protein coupled receptor that exclusively binds to CXCL12.[21]Some studies have demonstrated that the CXCL12–CXCR4 axis plays an important role in development, hematopoiesis, organ vascularization, and neuronal migration and is also involved in cancers and in organ-specific metastasis. [20] Because CXCR4 is overexpressed in several human cancers, the blockade of CXCR4–CXCL12 interactions has been extensively investigated as a potential cancer therapeutic. The lead CXCR4 antagonist, plerixafor (previously known as AMD3100), is the most clinically advanced compound. [22].

In particular, our peptides showed a stronger inhibitory capability on the adenilate cyclase compared to the AMD3100, i.e., . These results evidenced that the two short flexible peptides with no regular secondary structure could dynamically explore conformational ensembles.

The employment of an intrinsically disorder peptide as a biosensor or in molecular recognition could lie in the ability to control the transition between different structural states.[23] One opportunity to induce conformational peptide rearrangement could be offer by binding the bioactive hydrophilic sequences to an hydrophobic moiety .

In this contest it is well know that peptide amphiphiles (PAs) are a class of emergent molecules because they combine the structural features of amphiphilic surfactants with the functions of bioactive peptides[24-25] useful for biotechnological applications in drug delivery of pharmacological molecules.[26-29] In addition, PAs are able to self-assemble into a wide variety of nanostructures, such as spherical entities with a cylindrical geometry or rod like micelles, monolayers, bilayers, vesicles,

nanofibers. The PAs supramolecular aggregation is driven by simultaneous presence of hydrophilic moiety constituted by a bioactive peptide sequence eventually enriched with hydrophilic amino acids or by an ethoxilic linker and a hydrophobic moiety that usually consists of hydrocarbon groups, in a simple case, one or two alkyl chains long, typically containing more than ten carbons.

At light of these considerations we have designed lipophilic derivatives of Pep E and PepK bearing two stearic alkyl chains and/or ethoxylic spacer. Solution conformational preferences were analysed by CD, fluorescence, and DLS and biological tests were carried out by evaluating peptide abilities to interact with the CXCR4 receptor, inhibit CXCL12-induced migration and reduce cAMP in comparison to previous studies on other disordered linear peptides.

2. Materials and Methods

2.1 Reagents

All N-Fmoc-amino acid derivatives, 2-chlorotrityl chloride resin pre-loaded with Fmoc-Glu-OAllyl and Fmoc-Lys-OAllyl and coupling reagents were purchased from Calbiochem-Novabiochem (Laufelfingen, Switzerland, Iris Biotech GMBH or from Inbios (Napoli, Italy). Fmoc-21-amino-4,7,10,13,16,19-hexaoxaheneicosanoic acid (Fmoc-Ahoh-OH) was purchased from Neosystem (Strasbourg, France). N,N-Dioctadecylsuccinamic acid was synthesized according to published methods.[30]. All other chemicals were commercially available by Sigma-Aldrich (Milano, Italy) and were used as received unless otherwise stated.

Preparative RP-HPLCs were carried out on a LC8 Shimadzu HPLC system (Shimadzu Corporation, Kyoto, Japan) equipped with a UV lambda-Max Model 481 detector using Phenomenex (Torrance, CA) C4 (300 Å, 250 _ 21.20 mm, 5 m) columns. Elution solvents are H_2O/ 0.1% TFA (A) and CH3CN/0.1% TFA (B), from 20% to 95% over 20 minutes at 20 mL min^{-1} flow rate. Purity and identity were assessed by analytical LC–MS analyses by using Finnigan Surveyor MSQ single quadrupole electrospray ionization (Finnigan/Thermo Electron Corporation San Jose, CA), column C4-Phenomenex eluted with an H_2O/0.1% TFA (A) and CH3CN/0.1% TFA (B) from 20% to 95% over 20 minutes at 1 mL min^{-1} flow rate.. Mass spectra were carried out on a MALDI-TOF Voyager-DE Perseptive Biosystem (Framingham, MA).

2.2 Amphiphilic peptide synthesis

All amphiphilic peptides were synthesized following solid phase peptide standard protocol using Fmoc chemistry on 2-chlorotrityl chloride resin pre-loaded with the side chain of α Fmoc-Lys-OAllyl and α Fmoc-Glu-OAllyl, respectively (0.24 mmol scale). The sequence: H-YGECPC was manufactured on both resin. In particular the swelling and Fmoc-deprotection steps were performed

by treating the resin for 30 min in dichloromethane (DCM) and in dimethylformamide (DMF) /Piperidine (Pip)(80/20, v/v) solution, respectively.

For amino acid coupling washing steps after coupling and deprotection steps were performed by treating resin in DMF (3 x 1 min), MeOH (1 x 1 min), and DCM (3 x 1 min), successively. The resin was immersed and mixed in a DMF solution containing the Fmoc-protected amino acid (3 eq), O-benzotriazole-tetramethyl-uronium-hexafluoro-phosphate (HBTU) (3 eq), and diisopropylethylamine (DIEA) (6 eq) for 2 h at room temperature. The Fmoc ethoxylic spacer and lipophilic N,N-dioctadecyl succinic acid and were coupled, as previously described.[31]

Peptides were cleaved by immersing the resin in TFA/H_2O/triisopropylsilane (TIS) (30 mL, 95/2.5/2.5, v/v/v) for 3 h. Cleavage mixtures were filtered from the resin, peptides and crude products were precipitated by adding cold water, centrifuged, and decanted. Resulting white solids were dissolved in water and acetonitrile 4:1 v:v (10 mL) mixture and freeze-dried to give a white powder that was analyzed for purity and purified by preparative RP-HPLC, using the specified conditions eluted with the H_2O/0.1% TFA (A) and CH_3CN/0.1% TFA (B) mixture; a two steps gradient was used from 20% to 80% of B in 10 min and from 80% to 95% of B in 15 min with a flow rate of 20 mL/min. Purified peptides were characterized by LC-ESI-MS or by MALDI-TOF. Mass values found: 1443 a.m.u. [(C18)2 PepE]; 1442 a.m.u. [(C18)2 PepK];
1776 a.m.u. [(C18)2-L-PepK]; 1777 [(C18)2-L-PepE].

2.3. Preparation of the solutions

All solutions were prepared by dissolving the PAs lyophilized in aqueous solution. In details, $(C18)_2$-PepK and $(C18)_2$-L-PepK were dissolved in 10 mM TRIS buffer at pH 8.0; whereas $(C18)_2$-PepE and $(C18)_2$-L-PepE in 10 mM phosphate buffer (PBS) at pH 7.4. The pH-meter was calibrated with three standards at pH 4.00, pH 7.00 and pH 10.00. In most cases the samples to be measured were prepared from stock solutions. Concentrations of all solutions were determined by absorbance on a UV-Vis measurements carried out on Thermo Fisher Scientific Inc (Wilmington, Delaware

USA) Nanodrop 2000c spectrophotometer equipped with a 1.0 cm quartz cuvette (Hellma) using a molar absorptivity (ε) of 1390 M^{-1} cm^{-1} at λ = 275 nm, which correspond to the wavelength of the tyrosine residue

2.4. Fluorescence measurements

Values of critical micellar concentration (CMC) of peptide amphiphiles were obtained by fluorescence measurements. Fluorescence spectra were recorded at room temperature on a Jasco Model FP-750 spectrofluorophotometer in a 1.0 cm path length quartz cell. Equal excitation and emission bandwidths were used throughout the experiments, with a recording speed of 125 nm min-1 and automatic selection of the time constant. The CMC were measured by using 8-anilino-1-naphthalene sulfonic acid ammonium salt (ANS) as the fluorescent probe.[32] Small aliquots of peptide aqueous solution were added to a fixed volume (1.00 mL) of $2.0 \cdot 10^{-5}$ M ANS fluorophore directly in the quartz cell. CMC values were determined by linear least-squares fitting of the fluorescence emission at 480 nm, upon excitation at 350 nm versus the amphiphile concentration. Fluorescence emission spectra of all PAs were recorded at $5 \cdot 10^{-5}$ M concentration and 25°C exciting peptide solution at 275 nm.

2.5. Dynamic light scattering characterization

Dynamic light scattering (DLS) measurements were carried out using a Zetasizer Nano ZS (Malvern Instruments, Westborough, MA) that employs a 173° backscatter detector. Other instrumental settings are measurement position (mm): 4.65; attenuator: 8; temperature 25 °C; cell: disposable sizing cuvette. DLS samples were prepared at a final concentration of 2.0 10^{-4} M and centrifuged at room temperature at 13 000 rpm for 5 min.

2.6. CD measurements

Far-UV CD spectra were recorded from 260 to 190nm on a Jasco J-810 spectropolarimeter equipped with a NesLab RTE111 thermal controller unit using a 1 mm quartz cell at 25 °C. CD spectra of alkylated derivatives of PepE and PepK were carried out at $2.0 \cdot 10^{-4}$M concentration in 2.0 mM phosphate buffer and 10 mM TRIS, respectively. Other experimental settings were: scan speed, 10 nm min^{-1}; sensitivity, 50 mdeg; time constant, 16 s; bandwidth, 1 nm. Each spectrum was obtained averaging three scans, and by subtracting contributions from other species in solution and converting the signal to mean residue ellipticity in units of deg cm^2 $dmol^{-1}$ res^{-1}.

2.7. Binding assay

Binding between CXCR4 and four peptides was evaluated as previously described.[33] Briefly 2.5x105 CCRF–CEM cells were pre-incubated with 10 μM antagonist peptides in binding buffer (PBS 1x plus 0.2% BSA and 0.1% NaN3) for 1h at 37°C, 5% CO_2 and then labeled for 45 minutes with anti-CXCR4 PE-antibody (FAB170P, clone 12G5, R&D Systems, Minneapolis, MN, USA). Cells were washed in PBS and analyzed by FACS Canto II cytofluorimeter (Becton Dickinson Immunocytometry Systems, Mountain View, CA, USA).

2.8. Migration Assay

CCRF-CEM cells migration was assayed in 24-well Transwell chambers (Corning Inc., Corning, NY) using inserts with an 8-μm pore membrane. Membranes were precoated with collagen (human collagen type I/III) and fibronectin (20 mg/mL each). CCRF-CEM cells were placed in the upper chamber ($1x10^5$ cells/well) in RPMI containing 1% BSA (migration media). Cells were pre-incubated for 45 min with CXCR4 antagonist and allowed to migrate toward 100 ng/ml CXCL12 in the lower chamber. After 16 h incubation, migrated cells were collected from the lower chamber and counted. The migration index was defined as the ratio between migrating cells in the experimental group and migrated cells in the control group.

2.9 cAMP Assay

CCRF-CEM cells ($1x10^6$) were incubated in presence of four peptides or Plerixafor (known as AMD3100) being a CXCR4 antagonist that has provided proof of concept for inhibition of the pathway[34] at different concentrations (1 and 10μM) in combination with forskolin (F) (1 μM) for 20 min, followed by stimulation with CXCL12 (100ng/ml) for 10 min. Controls include cells stimulated with CXCL12 and forskolin or forskolin alone in absence of anti-CXCR4 inhibitors. Cells are harvested and lysed with 0.1M HCl and cAMP levels was assayed using a direct competitive enzyme immunoassay (BioVision Incorporated) according to manufacture instructions.

3. Results and Discussion

3.1 Design and synthesis

The linear peptide sequence YGECPC was designed to resemble an IDR (Intrinsically Disordered Region) by selecting amino acids with disorder propensity. Subsequently, PepK and PepE were obtained by adding at the C-termini of these peptides, a charged residues,i.e., either a lysine or a glutamic acid. Moreover, the carboxylic terminus of both peptides was further functionalized by an allylic group. This organic function could give the additional chance to further functionalize the molecules with an active ingredient or a contrast agent, for example through the olefin metathesis reaction. Peptide amphiphiles $(C18)_2$-PepK and $(C18)_2$-PepE, reported in Fig 1, were obtained by conjugating two aliphatic chains of eighteen carbon atoms at the N-terminus of both sequences. Two others PAs derivatives, indicated in Fig 1 as $(C18)_2$-L-PepK and $(C18)_2$-L-PepE, were obtained by inserting the AhOh exoethylene linker between the peptide sequence and the hydrophobic double-tail. This linker permits to increase the hydrophilicity of PAs without modifying their final charge, thus avoiding intra- and inter-electrostatic interactions. On the other hand, the insertion of this linker allows distancing the bioactive peptide from the hydrophobic shell of the aggregate. The relative short length of this moiety was chosen in order to avoid possible formation of an hydrophilic pocket that could hide the amino acid sequences. Finally, as well-

known from the literature,[35] the presence of PEG chains can increase the *in vivo* blood circulation of self-assembled PA aggregates. PAs were synthesized according to standard solid phase peptide synthesis protocols using 2-Cl-(Trt)-Cl preloaded resin as polymeric support. The ethoxylic moiety and the two alkyl chains at the N terminus of both the peptide sequences were added as previously reported.[32] After RP-HPLC purification, the products were identified by ESI-MS spectrometry.

3.2 Aggregates preparation and structural characterization

Supramolecular aggregates of pure PA derivatives were obtained by dissolving the lipophilic derivatives in 10 mM phosphate buffer (pH 7.4) and in 10 mM TRIS buffer (pH 8.0) for PepE and PepK amphiphiles, respectively. The choose of buffers to use for solubilizing monomers was dictated by their different propensity to aggregate. According to the pKa values of the glutamic acid and lysine residues, in neutral or slightly basic water solution all PAs display a net charge, negative in PepE and positive in PepK. Self-aggregation properties of the amphiphilic peptides were investigated by fluorescence spectroscopy and dynamic light scattering (DLS).

Critical micellar concentration (CMC) values were estimated by fluorescence spectroscopy by using ANS molecule as fluorescent probe. The fluorescence intensity of ANS is strongly dependent from the surrounding environment. ANS fluorophore emits only hydrophobic environment such as the hydrophobic core of a micellar aggregate, whereas does not emit in water solution. CMC values can be estimated from the graphical break-point, reported in Fig. 2, by plotting the fluorescence intensity of ANS in the emission maximum at 470 nm as a function of the PA derivative concentration. As reported in Table 1, CMC values are $7.0 \cdot 10^{-6}$ and $1.9 \cdot 10^{-5}$ mol kg^{-1} for $(C18)_2$-PepK and $(C18)_2$-L-PepK respectively. These values indicate that the introduction of the ethoxilic linker slight affects the formation of aggregates capable of sequestering the ANS, but also confirms the high stability of both the resulting aggregates. These values are in good agreement with those previously found for peptide amphiphiles containing the same hydrophobic moiety. On the other hand, the CMC values of both the amphiphilic derivatives of PepE are higher than the

corresponding PepK derivatives. This behaviour indicates a lower propensity to aggregate of PepE with respect to PepK, probably ascribable to the different interaction of the charges with the buffer salts. The hydrodynamic radii, R_H, the diffusion coefficients, D, and the polydispersity indexes, PDI, of samples, reported in Table 1, were obtained by the DLS technique. Measurements were performed at $\theta = 173°$ on self-assembled PAs at a concentration of $2 \cdot 10^{-4} mol \cdot Kg^{-1}$ in the previously reported buffer solution. All aggregate solutions show a monomodal distribution indicating the presence of one population of aggregates. At infinite dilution, R_H values can be evaluated using translational diffusion coefficients in the Stokes-Einstein equation. An increase of the radius of 20% can be observed for $(C18)_2$-PepE (100 nm) with respect to $(C18)_2$-PepK (80 nm) (see Fig. 3). Instead, only a slight difference in radii can be detected for the derivatives containing the spacer. Fluorescence measurements were carried out by exciting buffer solutions at 275 nm in order to evaluate eventually shifts of tyrosine fluorophore emission as a consequence of aromatic quenching on the surface of the aggregates. In Fig. 4 (upper panel) emission spectra for different PepE derivatives are compared. The behaviour of tyrosines does not show significant differences in intensity or in wavelength shift. These results seem to exclude interaction between among peptide side chains on aggregate the surfaces of aggregates. Similar results were obtained for PepK derivatives (Fig. 4 lower panel).

3.3 Circular Dichroism measurement

Conformational behaviour of PepE and PepK amphiphiles was investigated by circular dichroism (CD) spectroscopy. Lyophilized peptides were dissolved in the suitable weakly basic buffer (TRIS or PBS, see experimental section) at $2 \cdot 10^{-4} M$ concentration. Under these experimental conditions, PAs are well above CMC concentrations determined by fluorescence, and the presence of supramolecular aggregates in aqueous solution are favored. In fig. 5 CD spectra for PepE (panel A) and PepK (panel B) derivatives are reported together with their corresponding conformational preferences obtained by the CAPITO web-server. $(C18)_2$-PepE does not show any propensity to fold as indicated by CD spectra with a negative minimum around 198 nm typical of unordered

structures. In addition the presence of the ethoxilic spacer does not influence significantly the dichroic bent of the PA. This trend was also confirmed by CAPITO analysis[33] depicted in the lower panel of Fig. 5. Instead in the CD spectra of PepK derivatives (see Fig. 5), in addition to the minimum at 196 nm, two weak maxima at 208 and 232 nm are also detectable. These spectra suggest that the replacement of the glutamic acid with the lysine at the C-terminus, induces a rearrangement of the secondary structure with a partial folding of the IDP.

3.5 Functional characterization: Involvement into the CXCR4 receptor pathway (è questo il senso?)

To assess the inhibitory capability of four peptides, $(C18)_2$-PepK, $(C18)_2$-L-PepK, $(C18)_2$-PepE, and $(C18)_2$-L-PepE on the function of the activated CXCR4 receptor, we evaluated: i) peptide binding to the receptor through an indirect binding assay, ii) the reduction of cyclic adenosine monophosphate (cAMP) levels (?), and iii) the inhibition of CXCL12-induced migration. The results were compared to Plerixafor (known as AMD3100) that is a CXCR4 antagonist that has provided proof of concept for inhibition of the pathway (Dovremmo forse riarrangiare? Sembrano frasi identiche al manoscritto precedente). A few studies showed that mobilization with Plerixafor plus Granulocyte-colony stimulating factor (G-CSF), that is a glycoprotein that stimulates the bone marrow to produce granulocytes and stem cells, reduces the incidence of failure to collect the minimum number of CD34 stem cells necessary for autologous stem cell transplantation. Therefore, on the basis of these studies, Plerixafor in combination with G-CSF has FDA approval for hematopoietic stem cell mobilization in patients with non-Hodgkin lymphoma and multiple myeloma. [34]

The binding of the four peptides to CXCR4 was evaluated in CCRF-CEM using a CXCR4 12G5 monoclonal antibody as previously described.[37] As shown in fig. 6, the four peptides did not displace 12G5 antibody compared to AMD3100, which strongly reduced antibody binding to CXCR4. This results may be explained considering the different binding sites on AMD3100/12G5

antibody and the four highly flexible and disordered peptides. To evaluate peptide ability to inhibit the activation of CXCR4 mediated intracellular pathways, cAMP and migration assays were carried out as well.

In general, CXCR4 after the binding to CXCL12 promotes the conformational change of the Gα subunit and replacement of the bound GDP by guanine nucleotide triphosphate (GTP) (see Fig. 7A). This exchange triggers the further conformation changes within the Gα subunit, which allows the trimeric G protein to be released from the receptor, and to dissociate into the GTP-bound Gα subunit and Gβ/Gγ dimer (Figure 7A). Both the activated components interact with various effector proteins. In particular, Gα regulates adenylyl cyclase (AC) that serves as an effector enzyme that catalyzes 5′adenosine triphosphate into cAMP and thereby activates cAMP-dependent protein kinase, which regulates a host of other downstream effectors including MAPK signaling pathway (Fig.7A).[37] When a peptide antagonist, such our four peptides, binds to CXCR4, it inhibits this mechanism (Fig. 7B). In details, Fig. 8 shows that: i) $(C18)_2$-PepK, $(C18)_2$-L-PepK, $(C18)_2$-PepE, and $(C18)_2$-L-PepE have an average dose reduction of the cAMP level of about 66% at the lower dose (1μM) and 84% at the higher dose (10μM), ii) the better inhibitory capability is obtained for $(C18)_2$-L-PepK and $(C18)_2$-PepE at 10μM with values of 93% and 95%, respectively, and iii) AMD3100 has a lesser inhibitory capability on the adenilate cyclase compared to the four peptides.

Finally, the migration was performed (Fig. 9) and four peptides reduced CCRF-CEM cells migration towards CXCL12 although less effectively than AMD3100. In particular, $(C18)_2$-PepK reduces the migration index from 1.75 to 1.42 fold, $(C18)_2$-L-PepK from 1.75 to 1.40 fold, $(C18)_2$-PepE from 1.75 to 1.49 fold, and $(C18)_2$-L-PepE PepE from 1.75 to 1.46 fold whereas AMD3100 from 1.75 to 1.15 fold.

4. Conclusions

We have explored the possibility that intrinsically disordered peptides could be used as polar heads connected to alkyl chains to generate new molecular buildings for drug delivery vehicles in cells

over-expressing the CXCR4 receptor. In our previous investigations, two peptides were synthesized and studied by NMR and CD spectroscopy.[32] These compounds were indicated as PepE and PepK due to the presence of a glutamic acid or a lysine residue at the C-terminus. The absence of regular secondary structure elements confirmed the natively disordered nature of these peptides. In the present study, we have investigated the opportunity to induce conformational rearrangement by modifying these peptide sequences with two alkyl chains of eighteen carbon atoms each and with an ethoxylic linker. In fact, the insertion on peptides of aliphatic moieties pushes the resulting amphiphilic molecules to self-assemble in supramolecular aggregates. Once confined on the external surface of the aggregates, peptides could undergo a disorder-to-order transition.32 As demonstrated by fluorescence and DLS studies, all the designed PAs are able to self-assemble in large aggregates with an average diameter ranging between 80 and 110 nm, independently fromtheir charge and from the presence of the PEG spacer. All four PAs are strongly aggregated and poorly soluble at high concentrations (0.50–1.0 mM). These characteristics made them unsuitable for characterization performed by means of solution NMR techniques; even after changing the experimental conditions such as implementing buffers at different pH and temperatures. CMC values comply with those previously found for peptide amphiphiles containing the same hydrophobic moiety and they point out the high stability of the aggregates.[28,29] The higher CMC values of $(C18)_2$-PepE and $(C18)_2$-L-PepE indicate a lower propensity to aggregate if compared with the corresponding PepK derivatives. Also after the insertion of the hydrophobic portion, PepE keeps its tendency towards the disordered/unfolded state from the conformational point of view. Whereas both the PepK derivatives show a partial folding, with respect to the free peptide sequence. To the best of our knowledge, these compounds are the first CPC motif containing amphiphilic peptides which are able to aggregate. Comparing the four peptide derivatives with previously studied PepE and PepK sequences14, we can point out that (i) all peptides did not show an evident binding affinity for CXCR4; (ii) the best result about the migration assay was obtained for PepE

sequences; and (iii) the highest cAMP percentages were obtained for $(C18)_2$-LPepE-(93%) and $(C18)_2$-L-PepK (95%) at 10 mM (Fig. 6). According to these preliminary results, these PA systems could be further modified to realize new supramolecular peptide-based compounds, able to be encapsulated by living cells for biomedical industrial application.

Acknowledgements: Marian Vincenzi thanks the Università Italo Francese (UIF) for financial support, project C2-1, 2014/38723, cap. 6.01.1810 UIF.

Abbreviations

AhOh 21-Amino-4,7,10,13,16,19-hexaoxaheneicosanoic acid
ANS 8-Anilinonaphthalene-1-sulfonate
CMC Critical micellar concentration
CD Circular dichroism
DCM dichloromethane
DIPEA N,N-Diisopropylethylamine
DLS Dynamic light scattering
DMF N,N-Dimethylformamide
DPC n-Dodecyl phosphatidylcholine
EDTA Ethylenediaminetetraacetic acid
Fmoc 9-Fluorenylmethoxycarbonyl
HATU, 2-(7-Aza-1H-benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate
HOBt, 1-Hydroxy-1,2,3-benzotriazole
PyBOP Benzotriazol-1-yl-oxytripyrrolidinophosphonium hexafluorophosphate
RP-HPLC Reverse-phase high-pressure liquid chromatography
SPPS Solid phase peptide synthesis
TFA Trifluoroacetic acid
TIS Triisopropylsylane;

Figure Captions

Fig 1. Schematic representation of $(C18)_2$-PepK, $(C18)_2$-Pep E and $(C18)_2$-L1-PepK and $(C18)_2$-L1PepE. The peptide sequences are reported using the one-letter amino acid code

Fig. 2 Fluorescence intensity of the ANS fluorophore at 470 nm as a function of PAs concentration; data are multiplied by a scale factor for a better comparison. CMC values are established from graphical break points.

Fig. 3 DLS profiles of $(C18)_2$-PepK (A) and $(C18)_2$-PepE (B) peptide amphiphiles at a concentration of $2 \cdot 10^{-4}$ M.

Fig. 4 Fluorescence emission spectra recorded exciting buffer solutions at 275 nmfor Pep E derivatives (up) and Pep K derivatives (down).

Fig. 5 CD spectra of: A) $(C18)_2$-PepE and $(C18)_2$-L-PepE; and B) $(C18)_2$-PepK and $(C18)_2$-L-PepK. In the lower panel, the corresponding CAPITO analyses, in which CD values at λ=200 nm are plotted versus λ=222 nm.

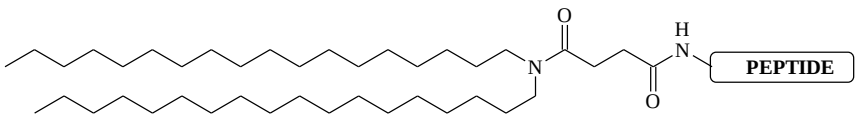
Fig. 6 Binding of peptides to CXCR4. The binding was evaluated indirectly through flow cytometry. In particular, histograms indicate the fluorescence percentage for CXCR4 antibody (CNT), AMD3100, $(C18)_2$-PepK, $(C18)_2$-L-PepK, $(C18)_2$-PepE, and $(C18)_2$-L-PepE.

Fig. 7 A schematic of the signaling pathway when CXCR4 binds to CXCL12 (A) or to a peptide antagonist (B) .

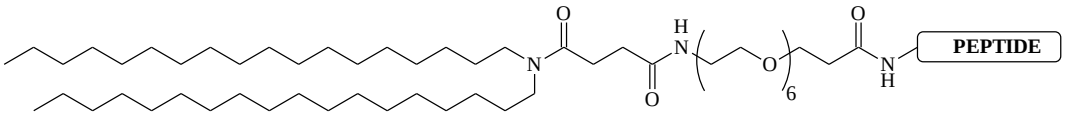
Fig. 8 Comparison of inhibition of cAMP modulation by AMD3100 and four peptides, $(C18)_2$-PepK, $(C18)_2$-L-PepK, $(C18)_2$-PepE, and $(C18)_2$-L-PepE, at 1μM and 10μM**.**

Fig. 9 Migration assays. We report the migration index relative to migration in presence of BSA alone, and of CXCL12, of AMD3100, and of four peptides, $(C18)_2$-PepK, $(C18)_2$-L-PepK, $(C18)_2$-PepE, and $(C18)_2$-L-PepE.

(C18)2-IDP



(C18)2-L-IDP



Peptide sequences:

PepK: Y-G-E-C-P-C-K-OAllyl

PepE: Y-G-E-C-P-C-E-OAllyl

Fig 1

Schematic representation of $(C18)_2$-PepK, $(C18)_2$-Pep E and $(C18)_2$-L-PepK and $(C18)_2$-L-PepE. The peptide sequences are reported using the one-letter amino acid code

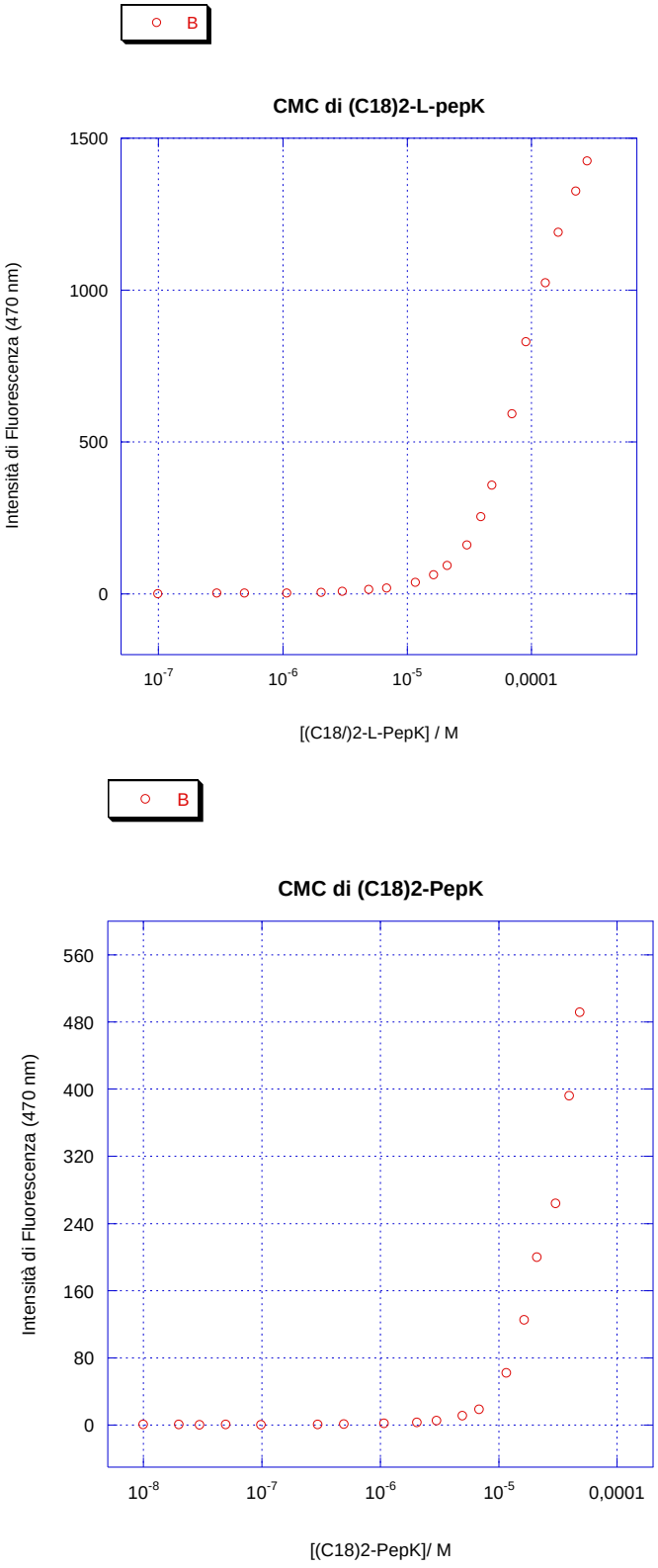


Fig. 2 Fluorescence intensity of the ANS fluorophore at 470 nm as a function of PAs concentration; data are multiplied by a scale factor for a better comparison. CMC values are established from graphical break points.

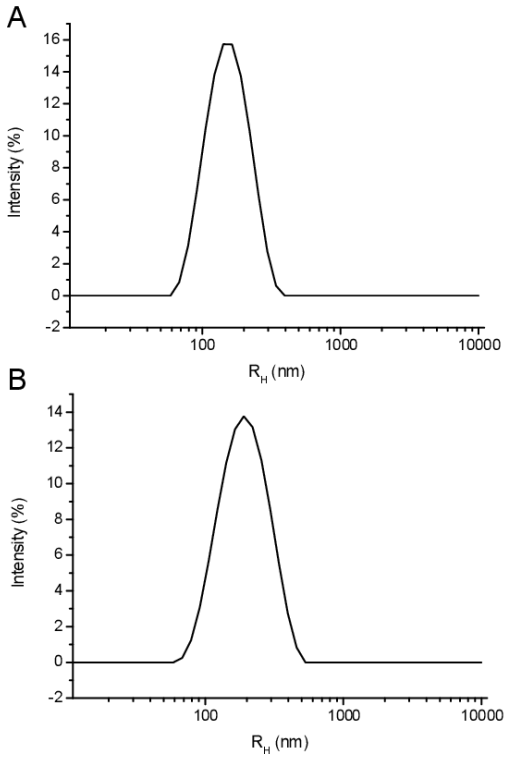


Figure 3

Fig. 3: DLS profiles of $(C18)_2$-PepK (A) and $(C18)_2$-PepE (B) peptide amphiphiles at a concentration of $2 \cdot 10^{-4}$ M.

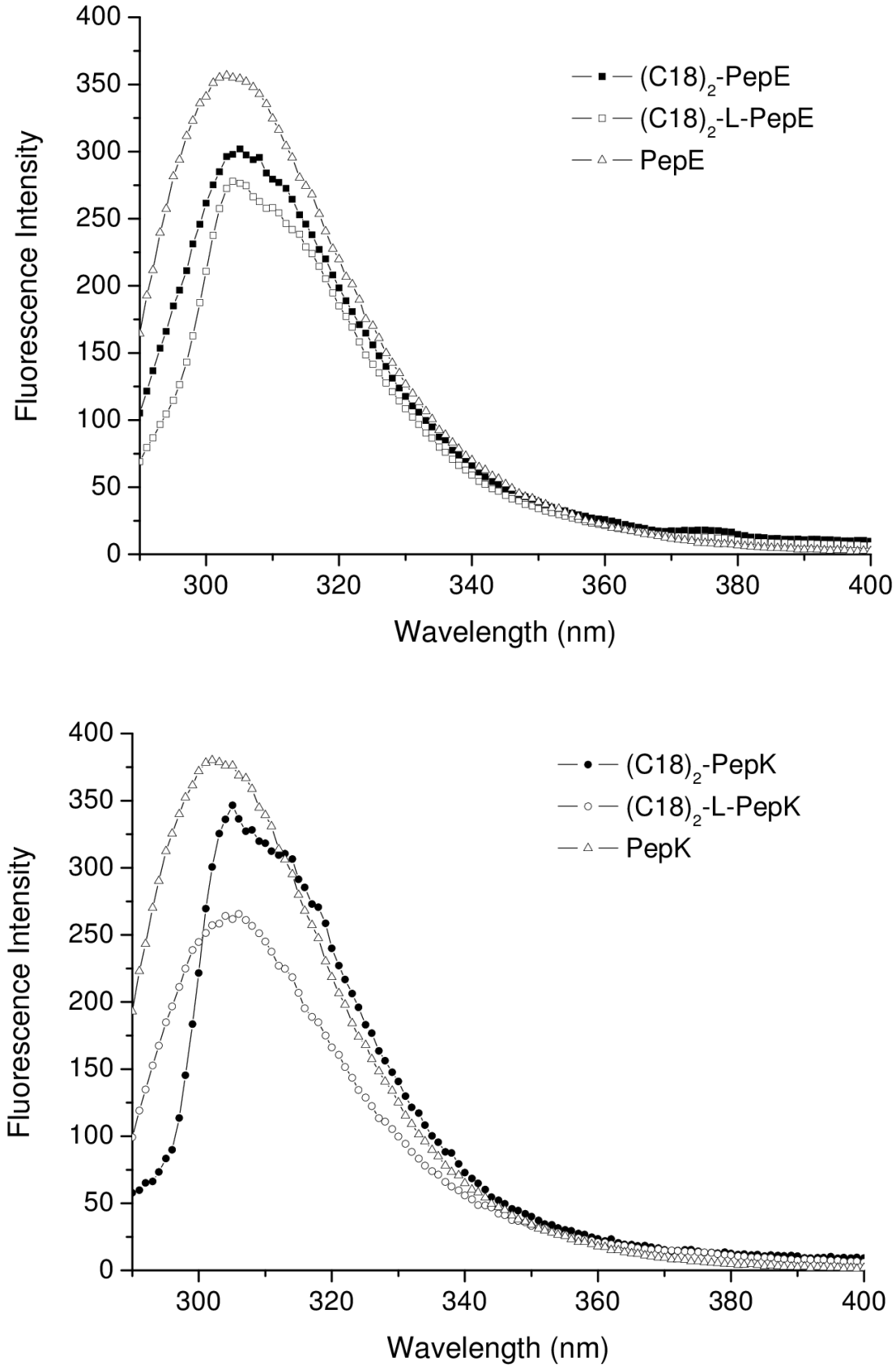


Fig. 4. Fluorescence emission spectra recorded exciting buffer solutions at 275 nmfor Pep E derivatives (up) and Pep K derivatives (down).

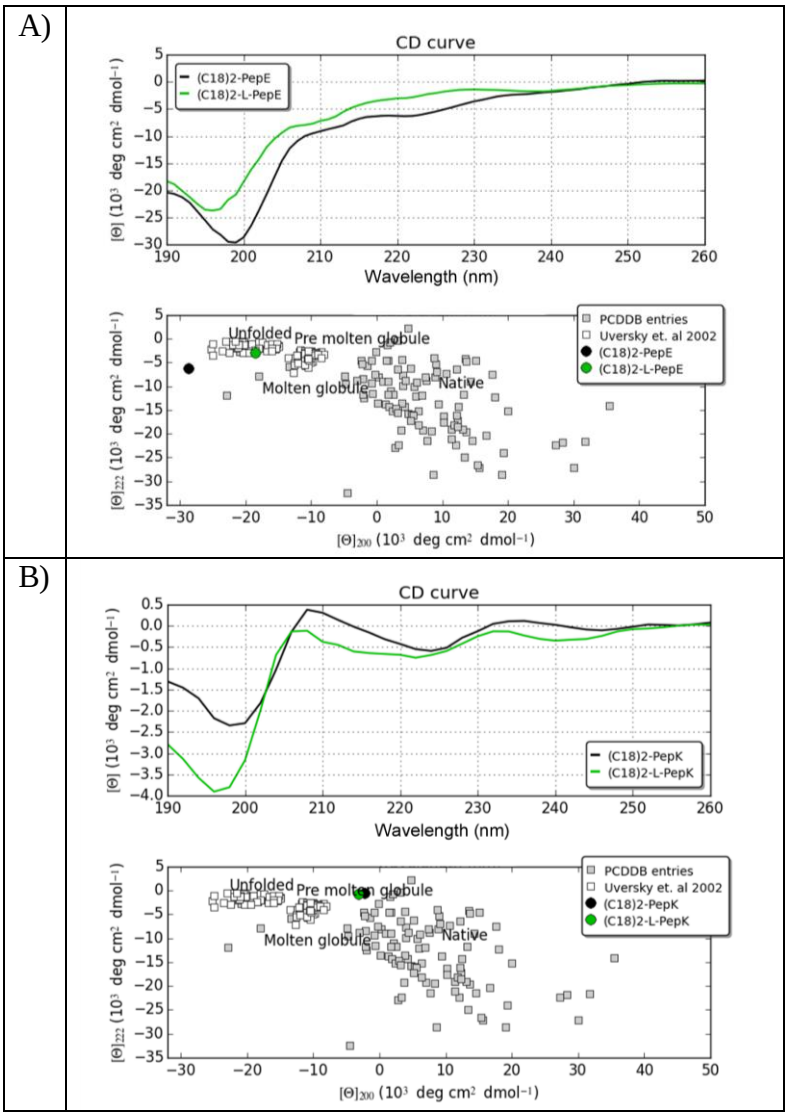


Fig. 5 CD spectra of: A) $(C18)_2$-PepE and $(C18)_2$-L-PepE; and B) $(C18)_2$-PepK and $(C18)_2$-L-PepK. In the lower panel, the corresponding CAPITO analyses, in which CD values at λ=200 nm are plotted versus λ=222 nm.

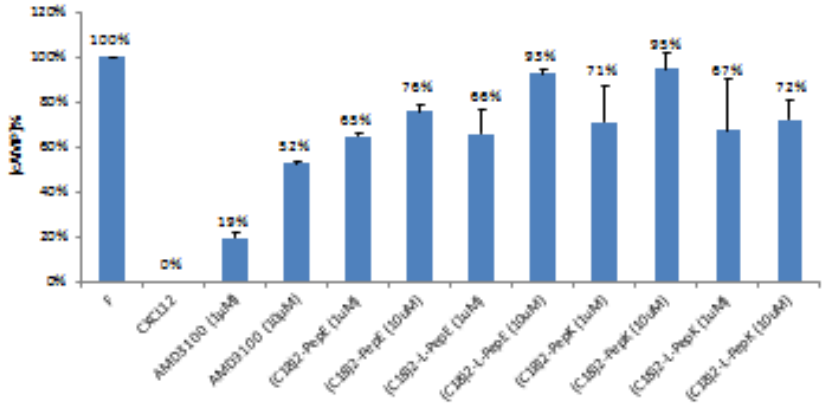
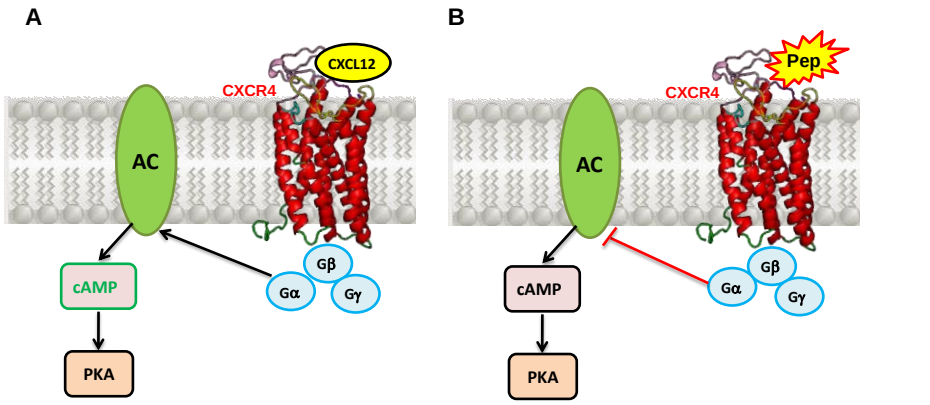


Fig. 6 Binding of peptides to CXCR4. The binding was evaluated indirectly through flow cytometry. In particular, histograms indicate the fluorescence percentage for CXCR4 antibody (CNT), AMD3100, $(C18)_2$-PepK, $(C18)_2$-L-PepK, $(C18)_2$-PepE, and $(C18)_2$-L-PepE.



CXCR4 after the binding to CXCL12 promotes the conformational change of the Gα subunit and replacement of the bound GDP by guanine nucleotide triphosphate (GTP). This exchange triggers the further conformation changes within the Gα subunit, which allows the trimeric G protein to be released from the receptor, and to dissociate into the GTP-bound Gα subunit and Gβ/Gγ dimer. Both the activated components interact with various effector proteins. In particular, Gα regulates adenylyl cyclase (AC) that serves as an effector enzyme that catalyzes 5′adenosine triphosphate into cyclic adenosine monophosphate (cAMP) and thereby activates cAMP-dependent protein kinase, which regulates a host of other downstream effectors including MAPK signaling pathway. When a peptide antagonist, such our four peptides, binds to CXCR4, it inhibits this mechanism.

Fig. 7: A schematic of the signaling pathway when CXCR4 binds to CXCL12 (A) or to a peptide antagonist (B) .

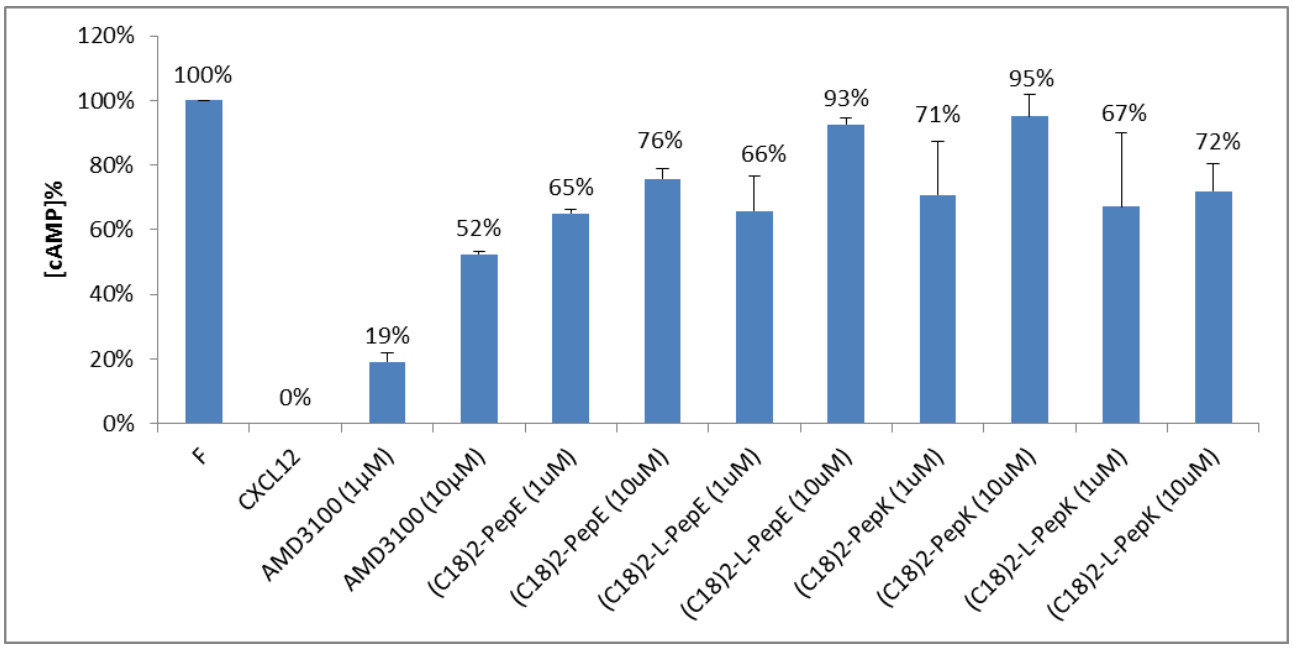


Fig 8 Comparison of inhibition of cAMP modulation by AMD3100 and four peptides, $(C18)_2$-PepK, $(C18)_2$-L-PepK, $(C18)_2$-PepE, and $(C18)_2$-L-PepE, at 1μM and 10μM**.**

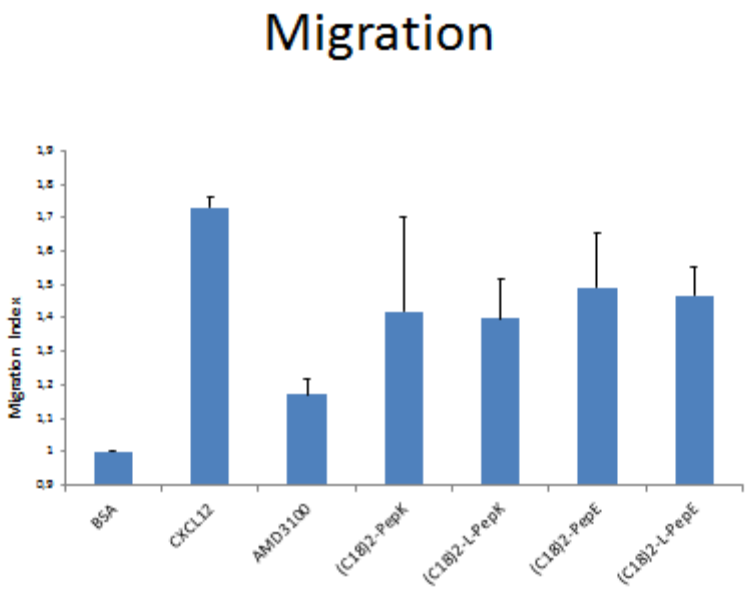


Fig. 9. Migration assays. We report the migration index relative to migration in presence of BSA alone, and of CXCL12, of AMD3100, and of four peptides, $(C18)_2$-PepK, $(C18)_2$-L-PepK, $(C18)_2$-PepE, and $(C18)_2$-L-PepE.

Table 1: Structural parameters for the aggregates obtained from fluorescence measurements (critical micellar concentration values) and dynamic light scattering measurements (diffusion coefficients, D, hydrodynamic radii, R_H, and polydispersity indexes).

Systems	**CMC (mol kg^{-1})**	**Rh (nm)**	**Dx10^{-12} (m^2 s^{-1})**	**PDI**
$(C18)_2$-PepK	$7.0 \cdot 10^{-6}$	80 ± 27	3.1	0.270
$(C18)_2$-L-PepK	$1.9 \cdot 10^{-5}$	82 ± 34	3.1	0.182
$(C18)_2$-PepE	$3.9 \cdot 10^{-5}$	112 ± 44	1.9	0.264
$(C18)_2$-L-PepE	$3.2 \cdot 10^{-5}$	101 ± 39	2.4	0.226

References

1. G. Wider, K. H. Lee and K. Wuthrich, *J. Mol. Biol*., 1982, 155,367–388.
2. L. Zetta, P. J. Hore and R. Kaptein, *Eur. J. Biochem*., 1983, 134, 371–376.
3. F. Naider, L. A. Jelicks, J. M. Becker and M. S. Broido, *Biopolymers*, 1989, 28, 487–497.
4. L. R. Brown, W. Braun, A. Kumar and K. Wuthrich, *Biophys. J*., 1982, 37, 319-.
5. J. K. Lakey, D.Baty andF.Pattus, *J. Mol. Biol*., 1991, 218, 639–653..
6. A. A. Kaspar, J.M. Reichert *Drug Discovery Today* 2013, 18(17/18), 807–817
7. K. Fosgerau, T. Hoffmann *Drug Discovery Today* 2015; ,20(1), 122–128
8. E. Benedetti, L. Aloj, G. Morelli, A. Accardo, R. Mansi, D. Tesauro Biodrugs 2004, 18, 279-295
9. M.F. Tweedle *Acc. Chem. Res.*, 2009, *42* (7), 958–968
10. V.J. Hruby, M. Cai *Annual Review of Pharmacology and Toxicology* 2013; 53, 557-580
11. S. Zhang *Materials Today* 2003, 6, 20-27
12. A. Lakshmanan, S. Zhang C.A.E. Hauser *Trends in Biotechnology*, 2012, 30, 155-165
13. EM Rezler, DR Khan, R Tu, M, Tirrell GB Fields. *Methods Mol Biol*. 2007;386:269-98.
14. J.C. Reubi *Endocr Rev* 2003, 24(4), 389-427.
15. R. Schweitzer-Stenner *Mol. BioSyst.*2012, 8, 122-133
16. V. N. Uversky, *Int J Biochem Cell Biol* 2011, 43, 1090-1103.
17. R. Raucci, G. Colonna, A. Giovane, G. Castello, S. Costantini, *Biochim. Biophys. Acta* 2014, 1844, 1868-1880.
18. J. Mittal, T. HyeonYoo, G. Georgiou T.T. Truskett, *J. Phys. Chem. B* 2013, 117, 118-124.
19. M. Vincenzi S. Costantini, S. Scala, D. Tesauro, A. Accardo, M. Leone, G. Colonna, Jean Guillon, L. Portella, A.M. Trotta, L. Ronga, F. Rossi *Int. J. Mol*. l. submitted
20. S. Costantini, R, Raucci. G. Colonna, F.A. Mercurio, A.M. Trotta, P. Ringhieri, M. Leone, F. Rossi, C. Pellegrino, G. Castello, S. Scala, *J.Pept. Sci*.2014, 20, 270-278]

21. L, Lasagni M, Francalanci F. Annunziato, E. Lazzeri, S. Giannini, L. Cosmi, C. Sagrinati, B. Mazzinghi, C. Orlando, E. Maggi, F. Marra, S. Romagnani, M. Serio P. Romagnani *J. Exp. Med*. 2003, 197, 1537–1549.
22. WT, Choi S. Duggineni, Y. Xu, Z, Huang *J. Med. Chem*. 2012, 55, 977–994.
23. S. Banta, Z. Megeed, M. Casali, K. Rege, M. L. Yarmush, *J. Nanosci. Nanotechnol*., 2007, 7, 387–401.
24. I. W. Hamley, *Soft Matter*, 2011, 7, 4122–4138
25. F. Versluis, H. R. Marsden, A. Kros, *Chem. Soc. Rev*., 2010, 39, 3434–3444..
26. Accardo, L.Aloj, M. Aurilio G. Morelli D. Tesauro,. *Int. J Nanomedicine* 2014, 9, 1537-1557
27. T.R. Pearce , K. Shroff , E. Kokkoli *Adv. Mater.* 2012, 24, 3803–3822.
28. A. Accardo, G.Salsano, A.Morisco, M.Aurilio, A. Parisi, F. Maione, C. Cicala, D. Tesauro, L. Aloj, G. De Rosa, G. Morelli, *Int. J. . Nanomed*.2012, *7* 2007–2017.
29. Y.Zhang, H.Zhang, X. Wang, J. Wang, X. Zhang, Q. Zhang *Biomaterials* 2012, 33 679-691
30. L. Schmitt and C. Dietrich, *J. Am. Chem. Soc.*, 1994, 116,8485–8491
31. A. Accardo, R.Mansi, A. Morisco, G. Mangiapia, L. Paduano, D. Tesauro, A. Radulescu, M. Aurilio, L. Aloj, C. Arra, G. Morelli, *Mol Biosys* 2010, 6, 878–887.
32. A. Accardo, D. Tesauro G Mangiapia, C. Pedone, G. Morelli *Pept. Sc. Biopolymers* 2007, 88, 115-121
33. Wiedemann, C.I., Bellstedt, P., Görlach M. *Bioinformatics* 2013, 29, 1750-1757
34. E. De Clercq, *Biochem. Pharmacol.* 2009, 77, 1655–1664.
35. L. Portella, R. Vitale, S. De Luca, C. D'Alterio, C. Ieranò, M. Napolitano, A. Riccio, M.N. Polimeno, L. Monfregola, A. Barbieri, A. Luciano, A. Ciarmiello, C. Arra, G. Castello, P. Amodeo, S. Scala, *PLoS One* 2013, 8(9), e74548.
36. Y. Matsumura, H. Maeda *Cancer Res*. 1986;46(12 Pt 1):6387–6392.
37. B.A. Teicher, S.P. Fricker *Clin. Cancer Res*.2010, 16, 2927-2931

38. A. Accardo, M. Leone, D. Tesauro**,** R. Aufiero, A. Bénarouche, J-F Cavalier, S. Longhi , F. Carriere, F. Rossi *Mol Biosys* 2013**,** 9, 1401--1410.

39.

Diminished Oligomerization in the Synthesis of New Anti-Angiogenic Cyclic Peptide Using Solution Instead of Solid-Phase Cyclization

Sandra Rubio,[1,2†] Jonathan Clarhaut,[3,4†] Elodie Péraudeau,[3,5] Marian Vincenzi,[1,2,6] Claire Soum,[1,2] Jean Guillon,[1,2] Sébastien Papot,[4] Luisa Ronga[1,2]

[1]*ARNA Laboratory, Université De Bordeaux, UFR Des Sciences Pharmaceutiques, Bordeaux Cedex F-33076, France*

[2]*ARNA Laboratory, INSERM U1212, 146 Rue Léo Saignat, Bordeaux Cedex 33076, France*

[3]*CHU De Poitiers, 2 Rue De La Milétrie, CS90577, Poitiers 86021, France*

[4]*Université De Poitiers, UMR-CNRS 7285, Institut De Chimie Des Milieux Et Des Matériaux De Poitiers, Groupe Systèmes Moléculaires Programmés, 4 Rue Michel Brunet, TSA 51106, Poitiers 86073, France*

[5]*Université De Poitiers, ERL-CNRS 7368, 1 Rue Georges Bonnet, TSA 51106, Poitiers 86073, France*

[6]*Department of Pharmacy, University of Naples "Federico II," and CIRPeB, via Mezzocannone 16, I-80134, Naples, Italy*

Received 16 November 2015; revised 13 January 2016; accepted 27 January 2016
Published online 00 Month 2016 in Wiley Online Library (wileyonlinelibrary.com). DOI 10.1002/bip.22814

ABSTRACT:

The design and synthesis of novel peptides that inhibit angiogenesis is an important area for anti-angiogenic drug development. Cyclic and small peptides present several advantages for therapeutic application, including stability, solubility, increased bio-availability and lack of immune response in the host cell. We describe here the synthesis and biological evaluations of a new cyclic peptide analog of CBO-P11: cyclo(RIKPHE), designated herein as CBO-P23M, a hexamer peptide encompassing residues 82 to 86 of VEGF which are involved in the interaction with VEGF receptor-2. CBO-P23M was prepared using in solution cyclization, therefore reducing the peptide cyclodimerization occurred during solid-phase cyclization. The cyclic dimer of CBO-P23M, which was obtained as the main side product during synthesis of the corresponding monomer, was also isolated and investigated. Both peptides markedly reduce VEGF-A-induced phosphorylation of VEGFR-2 and Erk1/2. Moreover, they exhibit anti-angiogenic activity in an in vitro morphogenesis study. Therefore CBO-P23M and CBO-P23M dimer appear as attractive candidates for the development of novel angiogenesis inhibitors for the treatment of cancer and other angiogenesis-related diseases. © *2016 Wiley Periodicals, Inc. Biopolymers (Pept Sci) 00: 000–000, 2016.*

Keywords: cyclic peptide; angiogenesis; CBO-P23M; CBO-P11; solid-supported cyclization; in solution cyclization; dimerization

This article was originally published online as an accepted preprint. The "Published Online" date corresponds to the preprint version. You can request a copy of any preprints from the past two calendar years by emailing the Biopolymers editorial office at biopolymers@wiley.com.

Correspondence to: Luisa Ronga, University of Bordeaux, UFR des Sciences Pharmaceutiques, U1212, ARNA Laboratory, F-33076 Bordeaux Cedex, France; e-mail: luisa.ronga@u-bordeaux.fr
Sandra Rubio and Jonathan Clarhaut contributed equally to this work
Contract grant sponsor: Agence Nationale de la Recherche (ANR, Programme Blanc-SIMI 7, ProTarget)
Contract grant sponsor: Università Italo Francese (UIF) (to M.V.)
Contract grant numbers: C2-1, 2014/38723, and 6.01.1810 UIF
Contract grant sponsor: Sport and Collection Association (to E.P.)

INTRODUCTION

Cancer is an angiogenesis-dependent disease[1,2] and blocking such angiogenesis is a well-established strategy to markedly reduce tumor growth and metastasis.[3]

J_ID: PEPS Customer A_ID: BIP22814 Cadmus Art: BIP22814 Ed. Ref. No.: BIP-PEP-2015-00113.R1 Date: 19-February-16 Stage: Page: 2

Among the numerous factors involved in angiogenesis, vascular endothelial growth factor (VEGF) is a potent mitogen for endothelial cells that plays a central role in angiogenesis by interaction with its specific receptors (VEGFR-1 or VEGFR-2) and subsequent activation of intracellular MAPK signaling pathways (e.g. Erk1/2).[4] Therefore, VEGF and its receptors VEGFR-1 and VEGFR-2 are prime targets for anti-angiogenic intervention which is thought to be one of the most promising strategies in cancer therapy. Several approaches have been employed to inhibit VEGF function. These include monoclonal neutralizing antibodies directed against VEGF and VEGF receptors, recombinant fusion proteins, recombinant soluble VEGF receptors, peptides that interfere with VEGF/VEGFR interactions and small organic molecules targeting the intracellular receptor kinase domain.[5]

A wide number of agents targeting both VEGF and its receptors have recently become standard treatments for different tumor types.[6] However, as tumor resistance to these agents may arise after a few months of treatment, new alternatives need to be investigated.

Within this context, our group has synthesized, characterized and evaluated as angiogenesis inhibitor several novel peptides derived from 79 to 93 VEGF sequence which is involved in the interaction with VEGFR-2.[7]

Among these analogs, the CBO-P11, a 17-amino acids cyclo peptide: cyclo(fPQIMRIKPHQGQHIGE), exhibits anti-angiogenic activity in ovo on the chicken chorioallantoic membrane (CAM) assay. Moreover, this peptide inhibits the growth of both established human intracranial and syngeneic glioma in mice.[8]

These results prompted us to investigate the anti-angiogenetic activity of a new, smaller CBO-P11 analog: cyclo(RIKPHE), designated herein CBO-P23M, which encompasses the shorter 82-86 VEGF sequence with the key residues Arg82, Lys84, and His86.[9] A glutamic acid was added to the C-terminal end of CBO-P23M in order to envisage a solid supported head to tail cyclization strategy and also for further functionalization. The extensive polymerization observed during the solid-supported cyclization led us to develop an alternative synthetic protocol employing cyclization in solution. This synthetic strategy allowed the preparation of CBO-P23M and its dimer as the main products which were isolated and tested in vitro as novel anti-angiogenesis agents.

Both our peptides markedly reduce VEGF-A-induced phosphorylation of VEGFR-2 and Erk1/2 and exhibit anti-angiogenic activity in an in vitro morphogenesis assay.

RESULTS AND DISCUSSION

Chemistry

In order to obtain CBO-P23M peptide, we decided to follow S1 the synthetic pathway depicted in Scheme 1. This synthesis is based on the three-dimensional protection strategy Fmoc/*t*Bu/OAllyl and the Glu side chain anchoring to the resin (Scheme 1, step i). This synthetic approach allows the orthogonal deprotection of the C- and N-*termini* for the subsequent on-resin head-to-tail cyclization (Scheme 1, step iv) [10].

While the solid-phase cyclization usually takes advantage of the pseudo-dilution phenomenon, in our synthesis, the on-resin cyclization favors the formation of the cyclodimerization product and only traces of the desired monomer are produced (Figure 1). Further polymerization by-products are also F1 observed: the cyclic trimer and tetramer. The attempts to isolate these products by RP-HPLC failed due to the small quantity of CBO-P23M formed and the co-elution of dimer, trimer and tetramer under the same HPLC peak at $t_R = 5.7$ min (Figure 1).

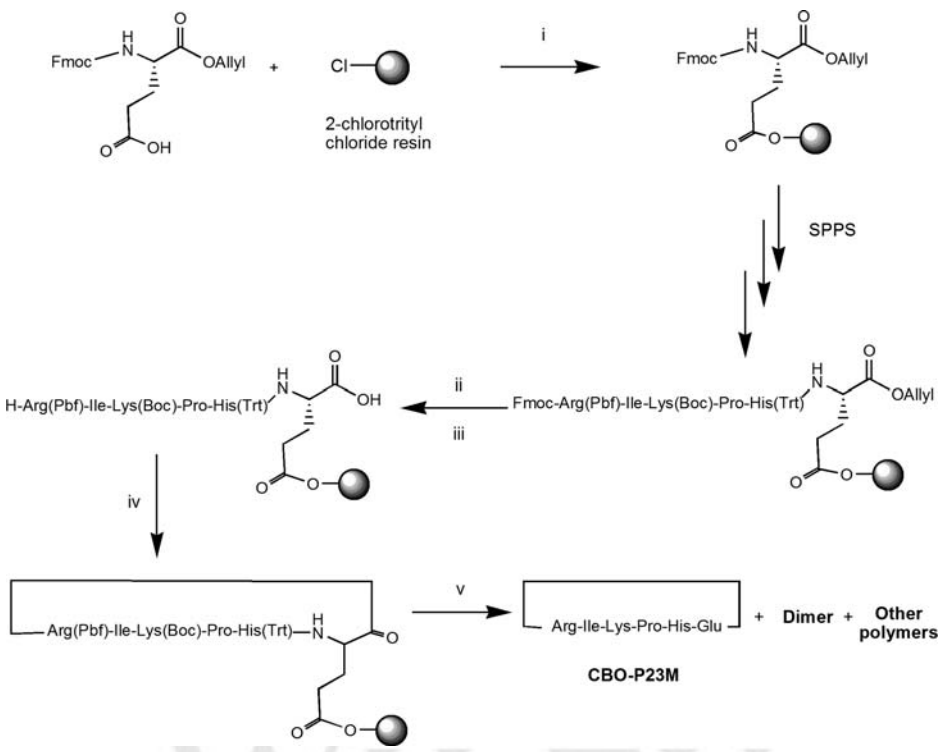
In order to avoid the formation of these side-products, the solid supported cyclization was investigated by modifying different parameters. Firstly, we reduced the cyclization time from 18 h to 1.5 h. Under these conditions only the monomeric, dimeric and trimeric linear peptides were obtained. Then, we tried the PyBOP/HOBt/DIEA coupling system. In this case, cyclo-dimerization and trimerization were mainly observed as for DIC/HOBt activation. Finally, we employed a low loaded resin (0.09 mmol/g) which was unfortunately not able to avoid the formation of cyclic polymers.

Under these circumstances, we next envisioned the synthesis of CBO-P23M peptide using the "in solution" cyclization of side-chain protected peptide (Scheme 2). For this purpose, we S2 started the elongation of our linear peptide from the proline residue in order to avoid epimerization of the C-terminal residue during peptide cyclization.[11] Then we performed the head to tail cyclization in solution under high diluted conditions and we obtained the expected CBO-P23M by reducing the cyclodimerization and the formation of the other polymerization side-products (Figure 2). The crude product was purified F2 and the CBO-P23M (Figure 3) and its dimer (Figure 4) isolated with (6% and 4% yield), respectively. F3 F4

Biological Evaluation

Because MAPK signaling pathway is one of the main target of VEGF-A, we analyzed, by Western Blotting, VEGF receptor 2 (VEGFR-2) and Erk1/2 phosphorylation by VEGF-A in absence or presence of CBO-P11,[8] CBO-P23M or CBO-P23M dimer. As shown in Figure 5, treatment of HUVECs with 20 F5 ng/mL of VEGF-A for 10 min strongly stimulates VEGFR-2 and Erk1/2 phosphorylation. Interestingly, 15 min of treatment (5 min of pretreatment and 10 min of co-treatment) with 20 μ*M* of CBO-P11, CBO-P23M, or CBO-P23M dimer markedly reduces VEGF-A-induced phosphorylation of VEGFR-2 and

ID: thangaraj.n Time: 22:40 I Path: //chenas03/Cenpro/ApplicationFiles/Journals/Wiley/PEPS/Vol00000/160006/Comp/APPFile/JW-PEPS160006



SCHEME 1 Synthesis of CBO-P23M by on resin cyclization; Reagents and conditions: (i) DIEA, DCM, 3 h; (ii) $PhSiH_3$, $Pd(PPh_3)_4$, 3 h; (iii) 20% piperidine in DMF, 2 × 20 min; (iv) Method a) DIC/HOBt, DIEA in DMF, 18 h; Method b) DIC/HOBt, DIEA in DMF, 1.5 h; Method c) PyBOP/HOBt, DIEA in DMF, 18 h (v) $TFA/H_2O/TIS$, 90/5/5, v/v/v, 3 h.

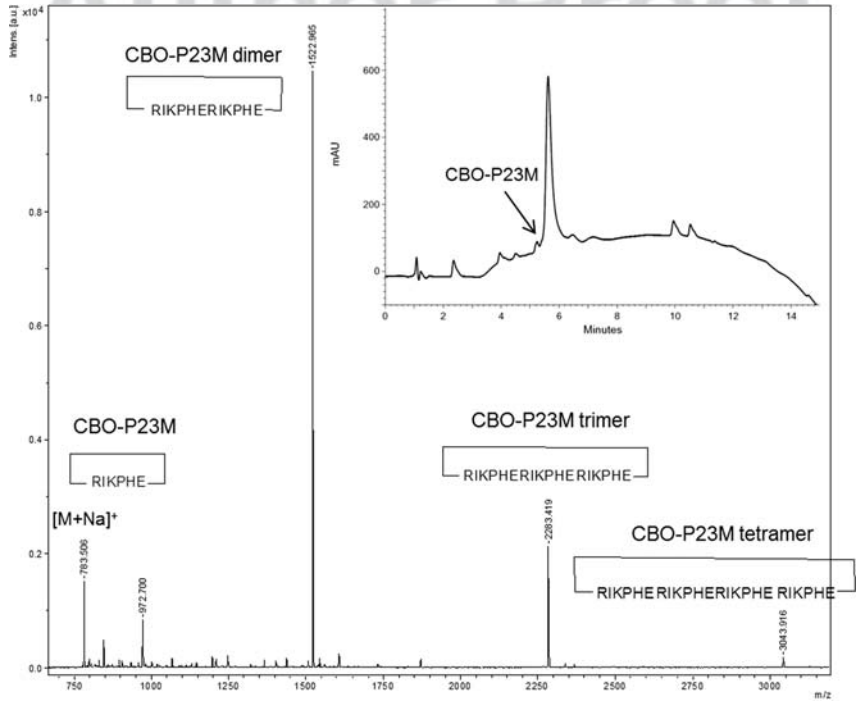
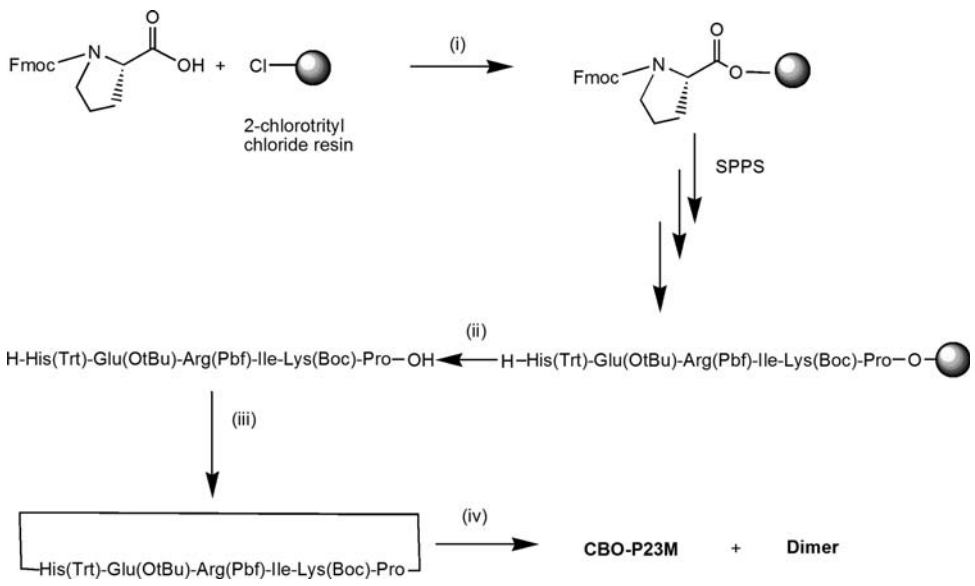


FIGURE 1 MALDI spectrum and HPLC profile of crude CBO-P23M obtained by on solid support cyclization (18 h reaction, DIC/HOBt as coupling system).

J_ID: PEPS Customer A_ID: BIP22814 Cadmus Art: BIP22814 Ed. Ref. No.: BIP-PEP-2015-00113.R1 Date: 19-February-16 Stage: Page: 4



SCHEME 2 Synthesis of CBO-P23M by in solution cyclization; reagents and conditions: (i) DIEA, DCM, 3 h, (ii) 1% TFA in DCM, 5 × 1 min, (iii) DIC/HOBt, DIEA in DCM, 18 h, (iv) TFA/H_2O/TIS, 90/5/5, v/v/v, 3 h.

Erk1/2. CBO-P23M and CBO-P23M dimer display almost the same inhibitor effect on VEGFR-2 phosphorylation (69.6% and 79.3%, respectively); higher than that observed with CBO-P11 (46.7%). On the other side, VEGF-induced phosphorylation of Erk1/2 is uniformly inhibited by CBO-P23M, its dimer and CBO-P11.

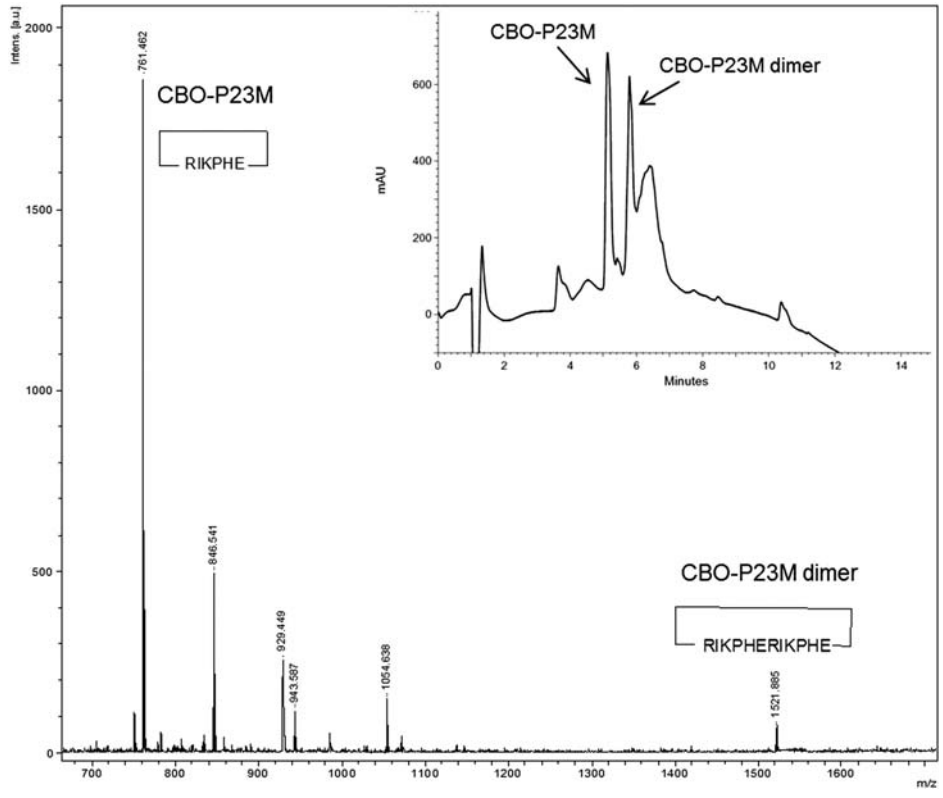


FIGURE 2 MALDI spectrum and HPLC profile of crude CBO-P23M obtained by in solution cyclization.

ID: thangaraj.n Time: 22:41 I Path: //chenas03/Cenpro/ApplicationFiles/Journals/Wiley/PEPS/Vol00000/160006/Comp/APPFile/JW-PEPS160006

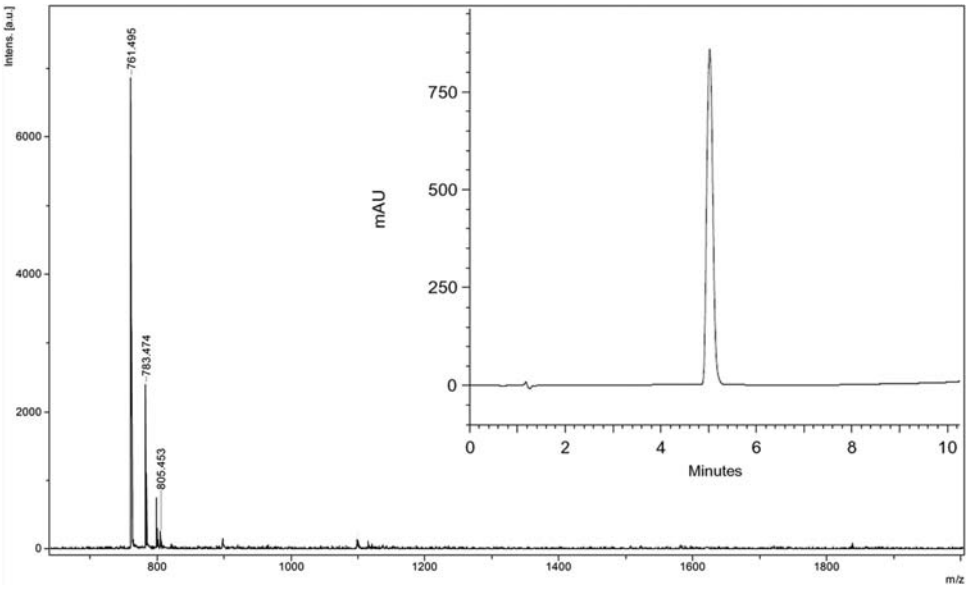


FIGURE 3 MALDI spectrum and HPLC profile of pure CBO-P23M.

Then, we investigated the capacity of HUVECs to form capillary-like structure by in vitro morphogenesis assay (Figure 6). This experiment is considered as the closest to in vivo angiogenesis[12] as it recapitulates several cellular events leading to vessel formation when HUVECs are cultured on extracellular matrix such as Matrigel®. Typical capillary-like structures were observed in basal condition (EGM-2 containing less than 2.5 ng/mL of VEGF-A) whereas, when cultured in presence of additional 20 ng/mL VEGF-A, the formation of the structures is stimulated by 44%. In contrast, treatments with 10 μ*M* of CBO-P11, CBO-P23M, or CBO-P23M dimer drastically reduced this network. Taken together, these in vitro morphogenesis results highlight an anti-angiogenic capacity of CBO-P23M and CBO-P23M dimer comparable to that of CBO-P11.

MATERIALS AND METHODS

Chemicals

2-Chlorotrityl chloride resin, Fmoc-protected amino acids (Fmoc-Arg(Pbf)-OH, Fmoc-Ile-OH, Fmoc-Lys(Boc)-OH, Fmoc-Pro-OH, Fmoc-Hys(Trt)-OH, Fmoc-Glu(tBu)-OH, Fmoc-Glu-OAllyl), *N*,*N*,*N*′,*N*′-tetramethyl-*O*-(1H-benzotriazol-1-yl)uranium hexafluorophosphate (HBTU), benzotriazol-1-yl-oxytripyrrolidinophosphonium hexafluorophosphate (PyBOP), Hydroxybenzotriazole (HOBt) and

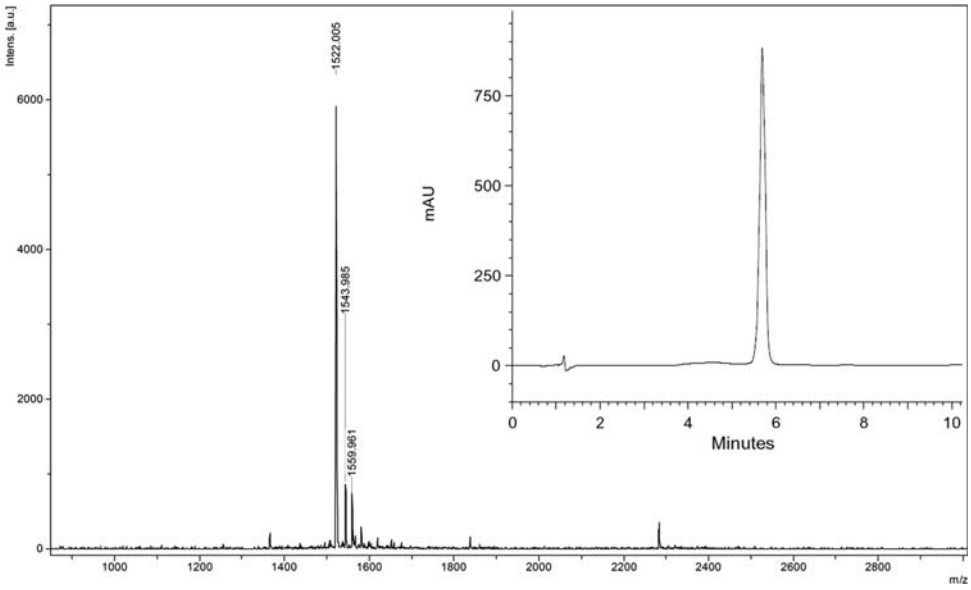


FIGURE 4 MALDI spectrum and HPLC profile of pure CBO-P23M dimer.

J_ID: PEPS Customer A_ID: BIP22814 Cadmus Art: BIP22814 Ed. Ref. No.: BIP-PEP-2015-00113.R1 Date: 19-February-16 Stage: Page: 6

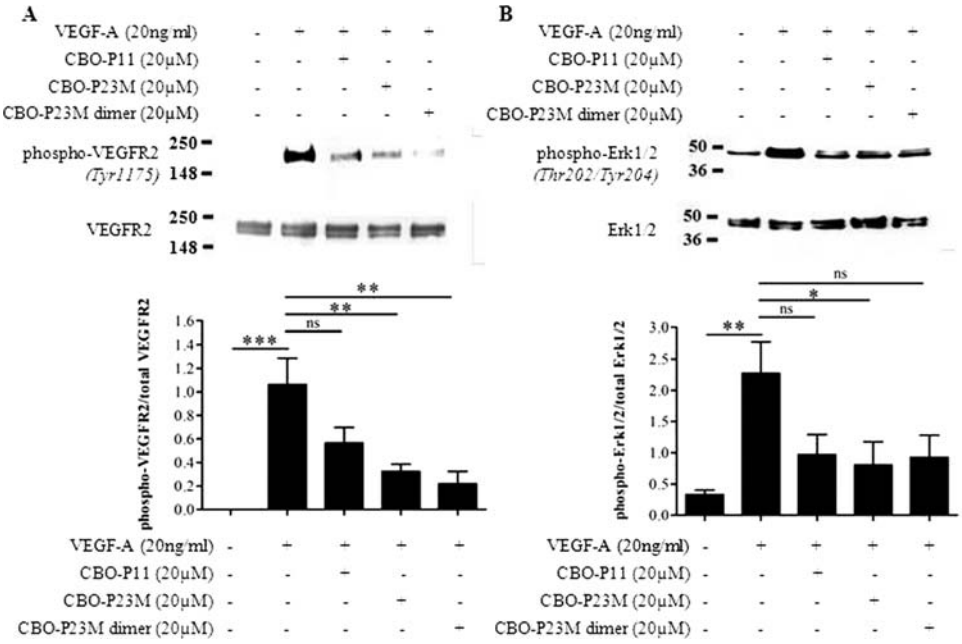


FIGURE 5 Inhibition of VEGF-A induced VEGFR-2 (A) and Erk1/2 (B) activation.

N,N′-diisopropylcarbodiimide (DIC) were obtained from Iris Biotech GMBH. All other chemicals were commercially available by Sigma-Aldrich and Acros.

Solid-Phase Peptide Synthesis General Experimental Procedures.

The syntheses were performed manually at 0.25 to 0.5 mmol scale using Fmoc/*t*Bu chemistry on 2-chlorotrityl chloride resin (1.6 mmol/g).

Resin Swelling. Swelling was performed by treating the resin for 20 min in dichloromethane (DCM).

First Amino Acid Coupling. After swelling of the resin, Fmoc-Pro-OH or Fmoc-Glu-OAllyl (2.3 eq) in DCM at 0.25*M* concentration and DIEA (6.9 eq) were added. The reaction mixture was left under mechanic stirring for 3 h at room temperature. The resin was left under vacuum overnight. Then the resin loading was determined from the Fmoc-release monitored by UV absorption at 301 nm.

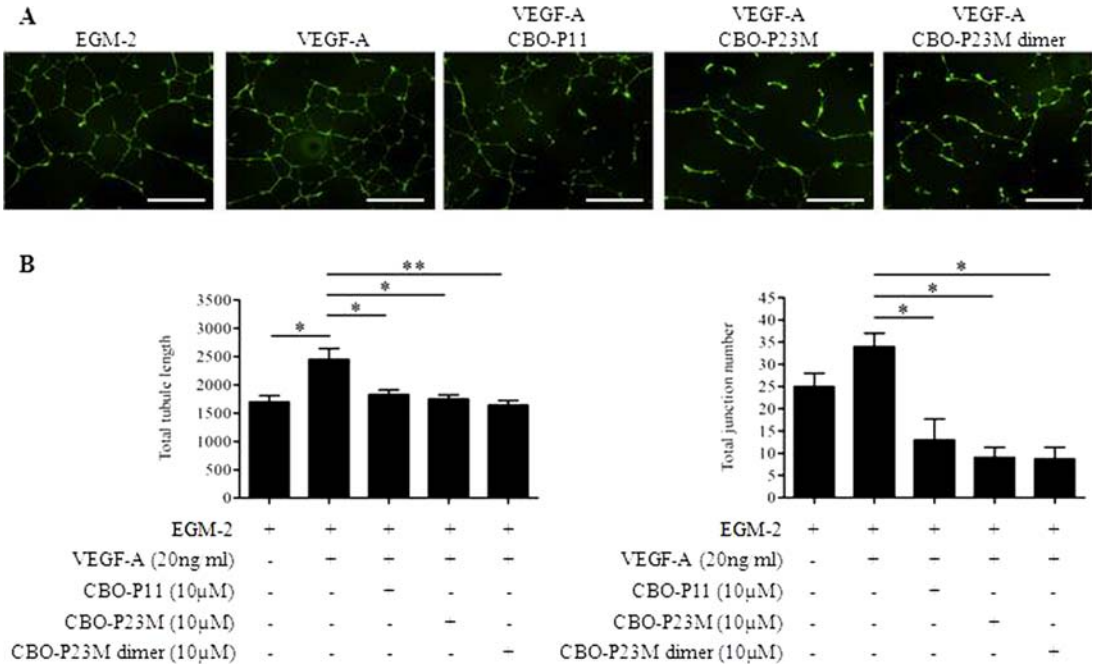


FIGURE 6 Inhibition of capillary-like network.

ID: thangaraj.n Time: 22:43 I Path: //chenas03/Cenpro/ApplicationFiles/Journals/Wiley/PEPS/Vol00000/160006/Comp/APPFile/JW-PEPS160006

J_ID: PEPS Customer A_ID: BIP22814 Cadmus Art: BIP22814 Ed. Ref. No.: BIP-PEP-2015-00113.R1 Date: 19-February-16 Stage: Page: 7

Fmoc-Deprotection. Fmoc-deprotection steps were performed by treating the resin 2 × 20 min in Piperidine (Pip)/DMF (80/20, v/v) solution.

Washing. Washing steps after coupling and deprotection steps were performed by treating the resin in dimethylformamide (DMF) (3 × 1 min), methanol (MeOH) (1 × 1 min), and DCM (3 × 1 min), successively.

Amino Acid Coupling. The resin was immersed in a DMF solution containing: Fmoc-protected amino acids (5 eq), HBTU (5 eq), and DIEA (10 eq) and left under mechanic stirring for 2 h.

Peptide Cyclization

On Resin. After peptide elongation, the Allyl-protection was removed from the peptide C-terminal carboxylate by treating the resin with $Pd(PPh_3)_4$ (0.35 eq) and $PhSiH_3$ (20 eq) in DCM, 3 h. The peptide was then cyclized on solid support by using DIC (5 eq) and HOBt (5 eq) in DMF overnight. Then the peptide was cleaved by immersing the resin in a trifluoroacetic acid (TFA)/H_2O/triisopropylsilane (TIS) (90/5/5, v/v/v) solution for 3 h. The cleavage cocktail was filtered from the resin and concentrated. The peptide was precipitated with ice-cold diethyl ether, centrifuged and decanted. Precipitation, centrifugation, and decantation operations were repeated twice. The resulting white solid was solubilized with H_2O/CH_3CN solution (50/50, v/v) and freeze-dried.

In Solution. After swelling of the resin, side-chain protected peptide was removed from the solid support by reacting the resin with a TFA/DCM (1/99, v/v) solution for 1 min, five times. Then the resin was washed with DCM (2 × 1 min). The cleavage and washing solutions were collected and evaporated. The product was solubilized with H_2O/CH_3CN solution (50/50, v/v) and freeze-dried.

The peptide was solubilized in DCM at 0.24 m*M* concentration and added dropwise (over 2 h) to a DCM solution containing HOBt (5 eq), DIC (5 eq), and DIEA (10 eq). This reaction mixture was left under magnetic stirring overnight and then evaporated.

Side-chains deprotection was performed by treating the peptide with a TFA/H_2O/TIS (90/5/5, v/v/v) solution for 3 h under magnetic stirring. Then the cleavage cocktail was concentrated and the peptide was precipitated with ice-cold diethyl ether, centrifuged and decanted. Precipitation, centrifugation, and decantation operations were repeated twice. The resulting white solid was solubilized with H_2O/CH_3CN solution (50/50, v/v) and freeze-dried.

For analysis, peptides were dissolved in a H_2O/CH_3CN/TFA solution (50:50:0.1, v/v/v) at 1 mg/mL concentration and analyzed by HPLC and Maldi.

RP-HPLC Analysis and Purification of Peptides

The HPLC characterization of products was performed by analytic reverse-phase HPLC using a VWR Hitachi instrument equipped with an L-2450 auto sampler, two L-2130 pumps, a Satisfaction RP18-AE column (5 μm, 250 × 4.6 mm) and a L-2450 diode array detector, at a flow rate of 0.8 mL/min. The compounds were purified by preparative reverse-phase HPLC using a VWR LaPrep system consisted of a P202 injector, two P110 pumps, a Satisfaction RP18-AB C18 column (5 μm, 250 × 20 mm) and a P314 UV detector, at a flow rate of 10 mL/min. The following eluents were used in a gradient mode: (A) 0.1% TFA in H_2O/CH_3CN (95/5) and (B) 0.1% TFA in CH_3CN/H_2O (95/5). Water was of Milli-Q quality and was obtained after filtration of distilled water through a Milli-Q® cartridge system. CH_3CN and TFA were of HPLC use quality. Degassing of solvents was performed using argon bubbling. Gradient used for the analytic and preparative RP-HPLC was respectively from 0% B to 100% B in 15 min and from 0% to 50% B in 50 min; UV detection at 214 nm. CH_3CN was evaporated and the aqueous solution was freeze-dried to give purified peptides as white solids.

Peptide Mass Analysis

Peptide identity was confirmed by mass spectrometry analyses performed on an Ultraflex III TOF/TOF system (Bruker Daltonics, Bremen, Germany), equipped with 200 Hz smartbeam laser (355 nm) and operating in reflectron positive ion mode. Mass spectra were acquired over the *m*/*z* range 300 to 5000 by accumulating data from 1000 laser shots for each spectrum. The instrumental conditions employed to analyze molecular species were the following: ion source 1: 25.08 kV; ion source 2: 21.98 kV, lens: 11.03 kV, pulsed ion extraction: 30 ns, reflector: 26.39 kV, reflector 2: 13.79 kV. Matrix suppression was activated by deflection mode: suppression up to 450 Da. Mass calibration was performed for each sample in range of 400 to 2000 Da with a peptide calibration mixture (8206195, Peptide Calibration Standard, Bruker Daltonics). The instrument was controlled using Bruker's flexControl 3.4 software and mass spectra were analyzed in Bruker's FlexAnalysis 3.4 software.

Cell Culture

Human umbilical vein endothelial cells (HUVECs, Lonza) were cultured in EGM-2 composed by endothelial cell basal medium-2 (EBM-2) and EGM-2 SingleQuot Kit Suppl. & Growth Factors (FBS, hydrocortisone, hFGF, VEGF, R3-IGF-1, ascorbic acid, hEGF, GA-1000 and heparin). HUVECs were used at early passages (1 to 8). Cells were kept at 37°C in a 5% CO_2 humidified incubator.

Western Blotting

HUVECs were plated in 12-well plates at a density of 5×10^4 cells/well in EGM-2 medium. After 3 days, cells were switched 16 h in basal medium EBM-2 before exposition to VEGF and peptides at the indicated concentrations and for the specified durations. Cells were collected by scraping and then lysed with Laemmli reducing buffer. Equivalent amounts of protein for each sample were analyzed by SDS-PAGE, transferred onto nitrocellulose membrane and probed with the following primary antibodies: anti-VEGFR2, anti-phospho-VEGFR Tyr1175, anti-Erk1/2 and anti-phospho-Erk1/2 Thr202/Tyr204 (Cell Signaling). Primary antibodies were detected using HRP-conjugated goat anti-mouse antibody (Dako) and Luminata HRP Substrate (Millipore). Densitometric analyses were performed in duplicate by 2 independent observers using ImageJ software. Statistical significance was determined by ANOVA test and Dunnet post-test. $^*P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$, ns: no statistical difference. Experiments were done three times.

ID: thangaraj.n Time: 22:44 I Path: //chenas03/Cenpro/ApplicationFiles/Journals/Wiley/PEPS/Vol00000/160006/Comp/APPFile/JW-PEPS160006

J_ID: PEPS Customer A_ID: BIP22814 Cadmus Art: BIP22814 Ed. Ref. No.: BIP-PEP-2015-00113.R1 Date: 19-February-16 Stage: Page: 8

Tube Formation Assay

Forty-eight-well plates were coated with 10 mg/mL Matrigel® (BD Biosciences) and kept at 37°C for 4 h to allow gel formation. HUVECs (5×10^4 cells/well) were then plated onto the Matrigel® in EGM-2 supplemented by VEGF-A (20 ng/mL) and investigated peptides (10 μ*M*). Sixteen hours later, networks were stained with calcein-AM (25 μ*M*, 30 min at 37°C) and fixed in PBS-2% formaldehyde (1 h at room temperature). The three-dimensional cell organization was photographed using MVX10 macroscope (Olympus). Capillary-like structures were quantified by automatic counting using the AngioQuant software. Images are representative of three independents experiments. Statistical significance was determined by ANOVA test and Dunnet post-test. $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$, ns: no statistical difference. Scale bar: 1 mm.

CONCLUSIONS

In summary, we designed the CBO-P23M as a novel cyclic peptide with potential anti-angiogenic activity. CBO-P23M is a cyclic hexamer peptide encompassing the 82 to 86 VEGF sequence with the key residues Arg82, Lys84, and His86 involved in the interaction with VEGFR-2. We demonstrated that CBO-P23M can be prepared using in solution cyclization, therefore overcoming the peptide cyclodimerization recorded during solid-phase cyclization

The anti-angiogenic activity of CBO-P23M was investigated together with that of its cyclic dimer (CBO-P23M dimer), obtained as the main side product during the synthesis of monomer. Our biological evaluations highlight an anti-angiogenic capacity of CBO-P23M and CBO-P23M dimer comparable to that of CBO-P11 peptide which was previously described by our group as promising anti-angiogenesis inhibitor. However, as CBO-P23M includes exclusively six amino acid residues, its synthesis appears much simpler than that of the 17-mer CBO-P11. Overall, these results make CBO-P23M an ideal candidate for the further development of new anti-angiogenic compounds.

ACKNOWLEDGMENTS

Luisa Ronga and Jean Guillon thank the "Plateforme Protéome" of the University of Bordeaux for the access to the Ultraflex III TOF/TOF mass spectrometer and the precious advice in using it.

REFERENCES

1. Folkman, J. N Engl J Med 1971, 285, 1182–1186.
2. Folkman, J. Nat Rev Drug Discov 2007, 6, 273–286.
3. Carmeliet, P.; Jain, R. K. Nature 2000, 407, 249–257.
4. Karkkainen, M. J.; Petrova, T. V. Oncogene 2000, 19, 5598–5605.
5. Tarallo, V.; De Falco, S. Int J Biochem Cell Biol 2015, 64, 185–189.
6. Zhao, Y.; Adjei, A. A. Oncologist 2015, 20, 660–673.
7. Betz, N.; Bikfalvi, A.; Deleris, G. Cyclic Peptides, Method for Preparing and Use as Angiogenesis Inhibitors or Activator. US Patent 2004/0092434 A1, May 13, 2004.
8. Zilberberg, L.; Shinkaruk, S.; Lequin, O.; Rousseau, B.; Hagedorn, M.; Costa, F.; Caronzolo, D.; Balke, M.; Canron, X.; Convert, O.; Laïn, G.; Gionnet, K.; Goncalvès, M.; Bayle, M.; Bello, L.; Chassaing, G.; Deleris, G.; Bikfalvi, A. J Biol Chem 2003, 278, 35564–35573.
9. Keyt, A. B.; Hung, V.; Nguyen, H. V.; Berleau, L. T.; Duarte, C. M.; Park, J.; Chen, H.; Ferrara, N. J Biol Chem 1996, 271, 5638–5646.
10. Trzeciak, A.; Bannwarth, W. Tetrahedron Lett 1992, 33, 4557–4560.
11. Lloyd-Williams, P; Albericio, F.; Giralt, E. Chemical Approaches to the Synthesis of Peptides and Proteins; CRC Press, 1997; pp 119. AQ4
12. Gallicchio, M.; Mitola, S.; Valdembri, D.; Fantozzi, R.; Varnum, B.; Avanzi, G. C.; Bussolino, F. Blood 2005, 105, 1970–1976.

ID: thangaraj.n Time: 22:44 I Path: //chenas03/Cenpro/ApplicationFiles/Journals/Wiley/PEPS/Vol00000/160006/Comp/APPFile/JW-PEPS160006

European Journal of Medicinal Chemistry xxx (2016) 1–14

Contents lists available at ScienceDirect

European Journal of Medicinal Chemistry

journal homepage: http://www.elsevier.com/locate/ejmech

Short communication

Synthesis and evaluation of the cytotoxic activity of novel ethyl 4-[4-(4-substitutedpiperidin-1-yl)]benzyl-phenylpyrrolo[1,2-*a*]quinoxaline-carboxylate derivatives in myeloid and lymphoid leukemia cell lines

Vanessa Desplat [a, b, 1], Marian Vincenzi [c, d, e, 1], Romain Lucas [c, d], Stéphane Moreau [c, d], Solène Savrimoutou [c, d], Noël Pinaud [f], Jordi Lesbordes [c, d], Elodie Peyrilles [c, d], Mathieu Marchivie [g], Sylvain Routier [h], Pascal Sonnet [i], Filomena Rossi [e], Luisa Ronga [c, d], Jean Guillon [c, d, *]

[a] *Univ. Bordeaux, UFR des Sciences Pharmaceutiques, Cellules souches hématopoïétiques normales et leucémiques, F-33076 Bordeaux cedex, France*
[b] *INSERM U1035, Cellules souches hématopoïétiques normales et leucémiques, F-33000 Bordeaux, France*
[c] *Univ. Bordeaux, UFR des Sciences Pharmaceutiques, ARNA Laboratory, F-33076 Bordeaux cedex, France*
[d] *INSERM U1212, UMR CNRS 5320, ARNA Laboratory, F-33000 Bordeaux, France*
[e] *Department of Pharmacy and CIRPeB, University of Naples "Federico II", Via Mezzocannone, 16 80134 Naples, Italy*
[f] *Univ. Bordeaux, ISM – CNRS UMR 5255, 351 cours de la Libération, F-33405 Talence cedex, France*
[g] *Univ. Bordeaux, ICMCB CNRS-UPR 9048, 87, Avenue du Docteur Schweitzer, F-33608 Pessac cedex, France*
[h] *Institut de Chimie Organique et analytique Univ. Orleans, CNRS UMR 7311, ICOA, BP 6759, rue de Chartres, 45067 Orléans cedex 2, France*
[i] *Université de Picardie Jules Verne, Laboratoire de Glycochimie, des Antimicrobiens et des Agroressouces, UMR CNRS 7378, UFR de Pharmacie, 1 rue des Louvels, F-80037 Amiens cedex 01, France*

ARTICLE INFO

Article history:
Received 29 January 2016
Received in revised form
17 February 2016
Accepted 18 February 2016
Available online xxx

Keywords:
Pyrrolo[1,2-a]quinoxaline
Synthesis
Cytotoxic activity
Cancer
Leukemia

ABSTRACT

Leukemia is the most common blood cancer, and its development starts at diverse points, leading to distinct subtypes that respond differently to therapy. This heterogeneity is rarely taken into account in therapies, so it is still essential to look for new specific drugs for leukemia subtypes or even for therapy-resistant cases. Among heterocyclic compounds that attracted a lot of attention because of its wide spread biological activities, the pyrrolo[1,2-*a*]quinoxaline heterocyclic framework has been identified as interesting scaffolds for antiproliferative activity against various human cancer cell lines. In the present study, novel ethyl 4-[4-(4-substitutedpiperidin-1-yl)]benzyl-phenylpyrrolo[1,2-*a*]quinoxaline-carboxylate derivatives **1a-l** have been designed and synthesized. Their cytotoxicities were evaluated against five different leukemia cell lines, including Jurkat and U266 (lymphoid cell lines), and K562, U937, HL60 (myeloid cell lines), as well as normal human peripheral blood mononuclear cells (PBMNCs). Then, apoptosis study was performed with the more interesting compounds. The new pyrrolo[1,2-*a*]quinoxaline series showed promising cytotoxic potential against all leukemia cell lines tested, and some compounds showed better results than the reference compound A6730. Some compounds, such as **1a**, **1e**, **1g** and **1h** are promising because of their high activity against leukemia and their low activity against normal hematopoietic cells. Structure-activity relationships of these new synthetic compounds **1a-l** are here also discussed.

1. Introduction

Acute leukemia is one of the most aggressive hematopoietic malignancies and is characterized by the abnormal proliferation of the immature cells and a premature block in lymphoid or myeloid differentiation. Adult acute leukemia have a poor prognosis due to a

* Corresponding author. INSERM U1212, UMR CNRS 5320, ARNA Laboratory, F-33000 Bordeaux, France.
E-mail address: jean.guillon@u-bordeaux.fr (J. Guillon).
[1] These authors contributed equally to the work.

http://dx.doi.org/10.1016/j.ejmech.2016.02.047
0223-5234/

large number of relapses. Thus, identifying and understanding the treatment-related resistance mechanisms is of major interest to improve the therapeutic strategy [1]. Therefore, there is an urgent need to find new therapeutics, which could led to the development of novel treatment strategies with less or minimal side effects.

Heterocyclic compounds attracted a lot of attention because of its wide spread biological activities. Among them, the pyrrolo[1,2-*a*]quinoxaline heterocyclic framework constitutes the basis of an important class of compounds possessing interesting biological activities. These compounds have been reported to serve as key intermediates for the assembly of several heterocycles including antipsychotic agents [2], anti-HIV agents [3], adenosine A_3 receptor modulators [4], antiparasitic agents [5–10], and antitumor agents [11–13]. In this last field, the discovery and development of novel therapeutic agents are one of the most important goals in medicinal chemistry. In this context, we have recently published three series (Series **A-C**) of new interesting substituted pyrrolo[1,2-*a*]quinoxalines (Fig. 1) endowed with good activity towards the human leukemia cells [14–16]. These antiproliferative pyrrolo[1,2-*a*]quinoxaline derivatives have been previously designed as novel structural analogues of compound A6730, a well-described Akt inhibitor that presents antiproliferative activity against different human leukemia cell lines [14–17]. Continuing our efforts in this field and considering the pharmacological activities of pyrroloquinoxalines on human leukemic cells, a new series (Series **D**) was designed and synthesized. Thus, by taking into accounts the best results obtained in series **B** (Fig. 1), we decided to use the JG576 and JG572 pyrrolo[1,2-*a*]quinoxaline moieties as a template for the design of new derivatives **1a-l** in which the pyrrole nucleus is substituted in different positions by a phenyl and an ester function (Series **D**, Fig. 1). In relation to our previous works, further pharmacomodulations on the piperidine core have been considered, such as the introduction of new substituted heterocyclic systems [14–16]. The antiproliferative profile of the obtained derivatives **1a-l** was then evaluated *in vitro* against a panel of myeloid (U937, HL60, K562) or lymphoid (Jurkat, U266) leukemic cell lines. Moreover, to determine their respective cytotoxicity, the new ethyl 4-[4-(4-substitutedpiperidin-1-yl)]benzyl-phenylpyrrolo[1,2-*a*]quinoxaline-carboxylate derivatives **1a-l** were tested on activated human peripheral blood mononuclear cells, and assessment of apoptosis was performed with the more interesting compounds. Structure-activity relationships of these new synthetic compounds **1a-l** are here discussed. Finally, we used simple computational programs to predict the drug-like characteristics through the calculated

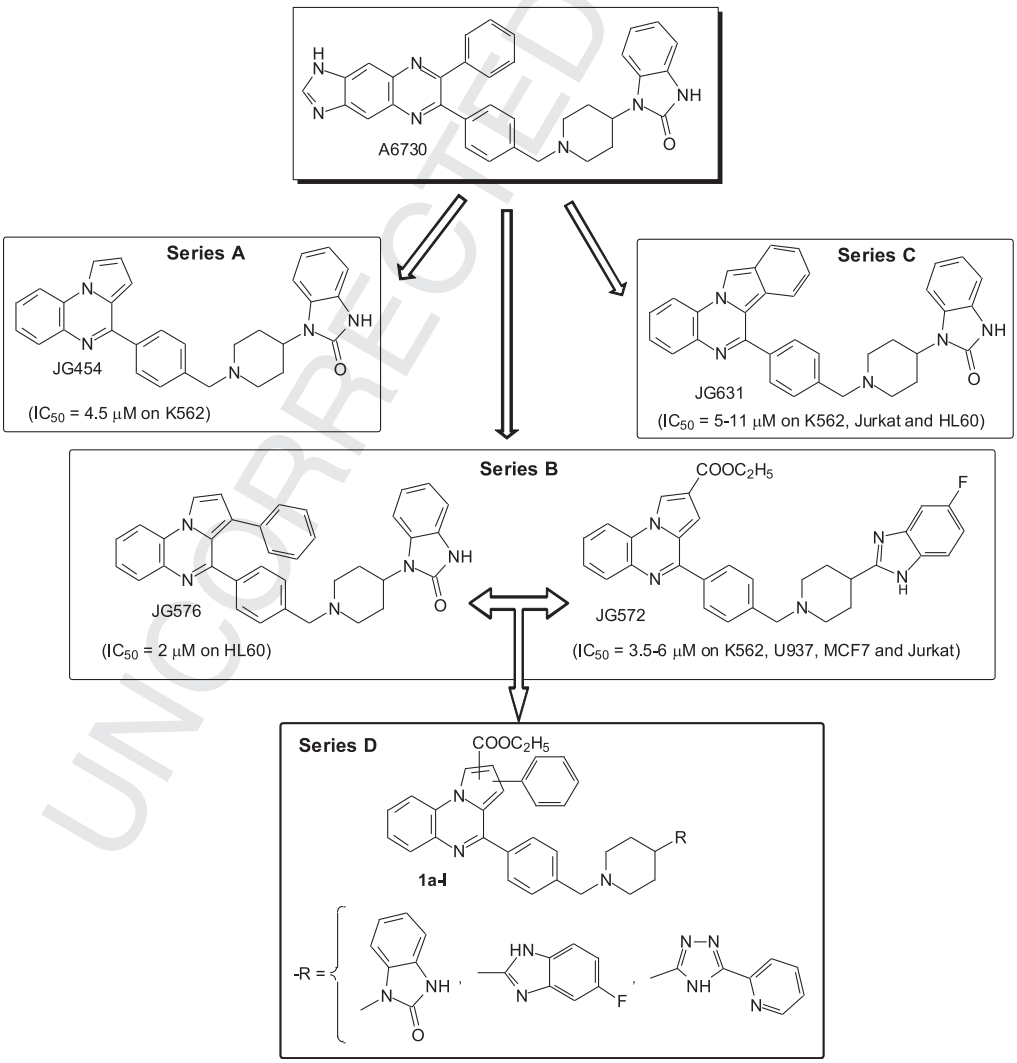


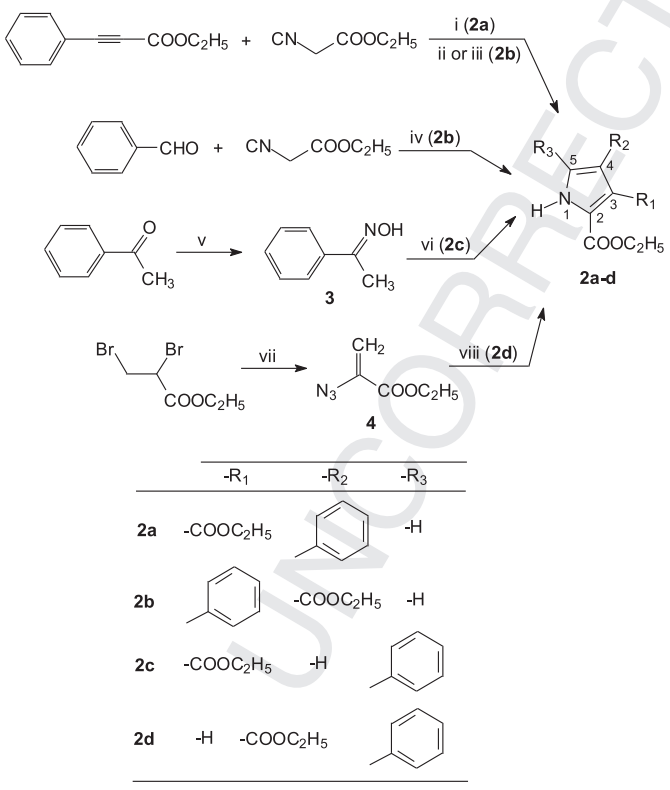
Fig. 1. Structure of bioactive compounds of previously described series A-C, and general structure of new synthesized substituted pyrrolo[1,2-a]quinoxaline derivatives **1a-l** (series D).

Please cite this article in press as: V. Desplat, et al., Synthesis and evaluation of the cytotoxic activity of novel ethyl 4-[4-(4-substitutedpiperidin-1-yl)]benzyl-phenylpyrrolo[1,2-*a*]quinoxaline-carboxylate derivatives in myeloid and lymphoid leukemia cell lines, European Journal of Medicinal Chemistry (2016), http://dx.doi.org/10.1016/j.ejmech.2016.02.047

physicochemical and toxicological properties of these new ethyl 4-[4-(4-substitutedpiperidin-1-yl)]benzyl-phenylpyrrolo[1,2-*a*]quinoxaline-carboxylate derivatives to determine their potential anti-leukemia activity.

2. Chemistry

All reported pyrrolo[1,2-*a*]quinoxaline derivatives **1a-l** were synthesized from various substituted phenyl-1*H*-pyrrole-dicarboxylic acid ethyl ester **2a-d** (Schemes 1 and 2). Different strategies using classical or microwaves heating were considered for the synthesis of the phenyl-1*H*-pyrrole-diesters **2a-d** in order to introduce the phenyl and ester functions on the pyrrole ring (Table 1). The synthesis of the diethyl 4-phenyl-1*H*-pyrrole-2,3-dicarboxylate **2a** has been accomplished by treatment of ethyl isocyanide on ethyl phenylpropiolate under 1,3-bis(diphenylphosphino)propane (dppp) catalysis via a formal [3 + 2] cyclo-addition (Scheme 1, Table 1) [18,19]. Various attempts were investigated for the preparation of the diethyl 3-phenyl-1*H*-pyrrole-2,4-dicarboxylate **2b**. At first, this pyrrole **2b** was prepared by silver-catalyzed cycloaddition of commercially available ethyl phenylpropiolate with ethyl isocyanoacetate in 1,4-dioxane or DMF at 80 °C (Scheme 1, Methods A-C, Table 1) [20,21]. The very low yield obtained (12%, Table 1) using microwaves heating led us to investigate other methodologies. The copper-catalyzed reaction of ethyl isocyanide with the electron-deficient alkyne, ethyl phenylpropiolate, gave the diethyl 3-phenyl-1*H*-pyrrole-2,4-dicarboxylate **2b** with 62% in dioxane at 100 °C using microwave heating (Table 1)



Scheme 1. Synthesis of phenyl-1H-pyrrole-diesters **2a-d**; Reagents and conditions: (i) dppp, dioxane, 100 °C; (ii) Method A: Ag_2CO_3, dioxane, 80 °C; Method B: 1) Ag_2CO_3, dioxane, 25 °C; 2) 80 °C; Method C: 1) Ag_2CO_3, NMP, 25 °C; 2) 80 °C; (iii) Method D: Cu_2O, phénanthroline, dioxane, 100 °C; (iv) Method E: DBU, THF, 50 °C; (v) H_2NOH, HCl, Pyridine, EtOH, reflux; (vi) 1) $H_5C_2OOC{-}C{\equiv}C{-}COOC_2H_5$, DABCO, toluene, 80 °C; 2) 170 °C, P; (vii) NaN_3, DMF, 65 °C; (viii) $C_6H_5{-}CO{-}CH_2{-}COOC_2H_5$, $Mn(OAc)_3$ $2H_2O$, 40 °C.

[1,2]. Pyrrole **2b** could be also synthesized by reaction of ethyl isocyanoacetate with benzaldehyde in the presence of 1,8-diazabicyclo[5.4.0]undec-7-ene (DBU) in THF [22,23]. The synthesis of the diethyl 5-phenyl-1*H*-pyrrole-2,3-dicarboxylate **2c** has been accomplished in two steps starting from commercially available acetophenone via its oxime **3** (Scheme 1). The acetophenone oxime **3**, synthesized by the reaction of hydroxylamine on acetophenone [7], was then reacted with diethyl acetylene in the presence of 1,4-diazabicyclo[2.2.2]octane (DABCO) to form pyrrole **2c** via a thermal rearrangement by Trofimov reaction [24,25]. Diethyl 5-phenyl-1*H*-pyrrole-2,4-dicarboxylate **2d** was prepared by manganese(III)-catalyzed formal [3 + 2] annulation of ethyl 2-azidoacrylate **4** and ethyl 3-oxo-3-phenylpropanoate [26]. The 2-azidoacrylate **4** was previously synthesized by treatment of ethyl 2,3-dibromopropanoate with three equivalents of sodium azide (NaN_3) in aprotic polar solvent such as dimethylformamide (DMF) [10,11]. A X-ray single crystal analysis was also performed on phenyl-1*H*-pyrrole-diesters **2b-d** in order to confirm the structures (Fig. 2). The preparation of *N*-aryl pyrroles **5a-d** were obtained by nucleophilic substitution of the various pyrrole-2-carboxylates **2a-d** with 2-fluoro-nitrobenzene using cesium carbonate as the base in refluxing DMF solution (Scheme 2) [15,16]. The preparation of **5a-d** was also performed under microwave irradiation. Reduction of the nitro moiety with iron in hot glacial acetic acid produced the spontaneous ring closure onto the ester in position 2 of the pyrrole moiety to afford the desired tricyclic pyrrolo[1,2-*a*]quinoxalines **6a-d** through a one-pot reduction-cyclization step [15,16]. The lactames **6a-d** were subsequently chlorodehydroxylated with phosphorous oxychloride, leading to the 4-chloroquinoxalines **7a-d**. 4-(Pyrrolo[1,2-*a*]quinoxalin-4-yl)benzaldehydes **8a-d** were easily prepared by a direct Suzuki-Miyaura cross-coupling reaction of 4-chloropyrroloquinoxalines **7a-d** with 4-formylphenylboronic acid performed in the presence of $Pd(PPh_3)_4$ as a catalyst, and in the presence of potassium carbonate used as the base [14–16]. The aldehydes **8a-d** were then engaged in a reductive amination with $NaBH_3CN$ and 4-(2-ketobenzimidazolin-1-yl)piperidine or 4-(5-fluorobenzimidazolin-2-yl)piperidine or 2-(3-piperidin-4-yl-1*H*-1,2,4-triazol-5-yl)pyridine to give the pyrrolo[1,2-*a*]quinoxalines **1a-l** [14–16]. The 3D spatial determinations of **1a**, **1d**, **1g** and **1j** were established by X-ray crystallography (Fig. 3), and confirmed the structures in the solid state as anticipated on the basis of IR and ^1H and ^{13}C NMR data.

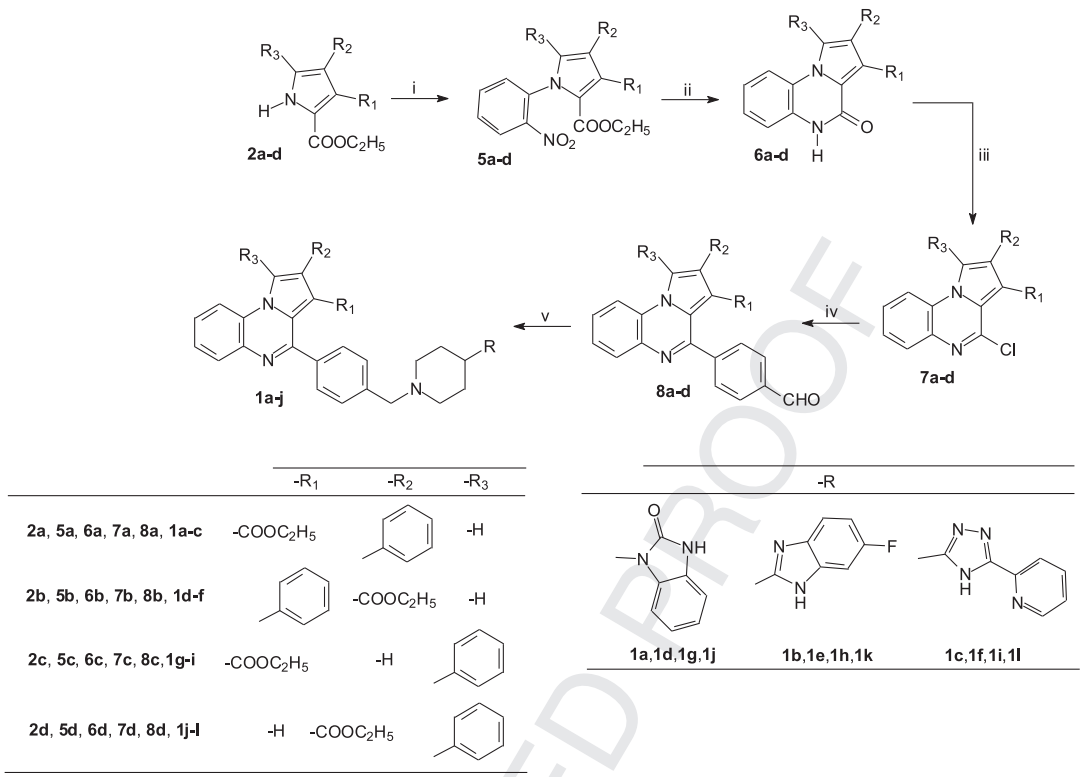
3. Biological activity

3.1. Cytotoxicity in leukemia cell lines

The twelve new compounds **1a-l** were tested in MTS assay for their *in vitro* antiproliferative activity against five human leukemic cell lines (U937, K562, Jurkat, U266 and HL60). Compound A6730 (Fig. 1) was used in these tests as the reference standard drug. The results are summarized in Table 2. In addition, compound LY-294002, which showed antiproliferative activity against the HL60, U937 and K562 cell lines [27–30], was also applied as a referential cytotoxic agent.

The pyrrolo[1,2-*a*]quinoxalines **1c**, **1f**, **1g**, **1h, 1i, 1k** and **1l** were found the most antiproliferative compounds on the growth of human myeloid U937 cell line with IC_{50} of 3–11 μM. The two derivatives **1h** and **1k** with IC_{50} of 4 and 3 μM respectively, showed a better activity in comparison with the reference compound A6730 ($IC_{50} = 8$ μM). These two compounds **1h** and **1k** were substituted in position 4 by a benzylpiperidinyl fluorobenzimidazole group and in position 1 by a phenyl. In general, the substitution by a phenyl on position 1 of the pyrrole moiety led to active derivatives (**1g-l** excepted **1j**). Interestingly, **1c**, **1f, 1i** and **1l** were substituted by a

Please cite this article in press as: V. Desplat, et al., Synthesis and evaluation of the cytotoxic activity of novel ethyl 4-[4-(4-substitutedpiperidin-1-yl)]benzyl-phenylpyrrolo[1,2-*a*]quinoxaline-carboxylate derivatives in myeloid and lymphoid leukemia cell lines, European Journal of Medicinal Chemistry (2016), http://dx.doi.org/10.1016/j.ejmech.2016.02.047



Scheme 2. Synthesis of pyrrolo[1,2-a]quinoxalines **1a-l**; Reagents and conditions: (i) 2-fluoro-nitrobenzene, Cs_2CO_3, DMF, Δ; (ii) Fe, CH_3COOH, Δ; (iii) $POCl_3$, Δ; (iv) $OHC{-}C_6H_4{-}B(OH)_2$, $Pd[P(C_6H_5)_3]_4$, K_2CO_3, toluene, EtOH, Δ; (v) 4-(2-ketobenzimidazolin-1-yl)piperidine or 4-(5-fluorobenzimidazolin-2-yl)piperidine or 2-(3-piperidin-4-yl-1H-1,2,4-triazol-5-yl]pyridine, $NaBH_3CN$, MeOH, Δ.

Table 1
Synthesis of phenyl-1*H*-pyrrole-diester **2a-d** under standard reaction conditions.

	Reagents and conditions	Time	Yield (%)
2a	(i) dppp, dioxane, 100 °C	2 h	87[a]
2b	(ii) *Method A*: Ag_2CO_3, dioxane, 80 °C	1 h	9[a]
	(ii) *Method B*: 1) Ag_2CO_3, dioxane, 25 °C; 2) 80 °C	1) 5 min 2) 30 min	10[a]
	(ii) *Method C:* 1) Ag_2CO_3, NMP, 25 °C; 2) 80 °C	1) 5 min 2) 30 min	12[b]
	(iii) *Method D:* Cu_2O, phenanthroline, dioxane, 100 °C	40 min	62[b]
	(iv) *Method E:* DBU, THF, 50 °C	1 h	35[b]
2c	(vi) 1) $H_5C_2OOC{-}C{\equiv}C{-}COOC_2H_5$, DABCO, toluene, 80 °C; 2) 170 °C, P	1) 6 min 2) 45 min	28[b]
2d	(viii) $C_6H_5{-}CO{-}CH_2{-}COOC_2H_5$, $Mn(OAc)_3$ $2H_2O$, 40 °C	2 h	42[a]

[a] Conventional heating.
[b] Microwaves heating (200 W).

benzylpiperidinyl triazolylpyridine moiety in position 4 of the pyrrolo[1,2-*a*]quinoxaline core, All other compounds derived from the incorporation of the benzylpiperidinyl benzimidazolone moiety (compounds **1a, 1d** and **1j**), which was present in the reference compound A6730, into the 4-position of the heterocyclic pyrroloquinoxaline ring were found inactive on the U937 cell line in comparison with their benzylpiperidinyl fluorobenzimidazole or benzylpiperidinyl triazolylpyridine analogues, excepted **1g** that was found active (IC_{50} = 9 μM). From a SAR point of view, these preliminary biological results on U937 cell line enlightened the importance of the substitution at C-4 position of the pyrroloquinoxaline scaffold by a benzylpiperidinyl fluorobenzimidazole group, and also the need of a phenyl functionalisation in position 1 of the pyrrole ring (compounds **1h** and **1k**).

The antiproliferative potencies of these new derivatives **1a-l** were also examined towards the human myeloid leukaemia cell lines K562 and HL60.

Among the twelve compounds tested for antiproliferative activities on K562 cell line, the five pyrrolo[1,2-*a*]quinoxalines **1a**, **1b**, **1e**, **1k** and **1l** were found the most active compounds with an IC_{50} of 3–4 μM. All the other pyrroloquinoxalines **1** also showed significant antiproliferative activity with IC_{50} ranging from 7 to 14 μM, better than the one found for the reference compound A6730 (IC_{50} = 17 μM). Moreover, in terms of structure-activity relationships discussion, it could be also noticed that the four quinoxalines **1b**, **1e**, **1h** and **1k**, bearing the benzylpiperidinyl fluorobenzimidazole moiety in their 4-position, were always found the most active compounds with IC_{50} of 3–7 μM in each subseries diversely substituted by a phenyl and an ester on the pyrrole ring.

Against the HL60 human acute promyeloid leukemia cell line,

Please cite this article in press as: V. Desplat, et al., Synthesis and evaluation of the cytotoxic activity of novel ethyl 4-[4-(4-substitutedpiperidin-1-yl)]benzyl-phenylpyrrolo[1,2-*a*]quinoxaline-carboxylate derivatives in myeloid and lymphoid leukemia cell lines, European Journal of Medicinal Chemistry (2016), http://dx.doi.org/10.1016/j.ejmech.2016.02.047

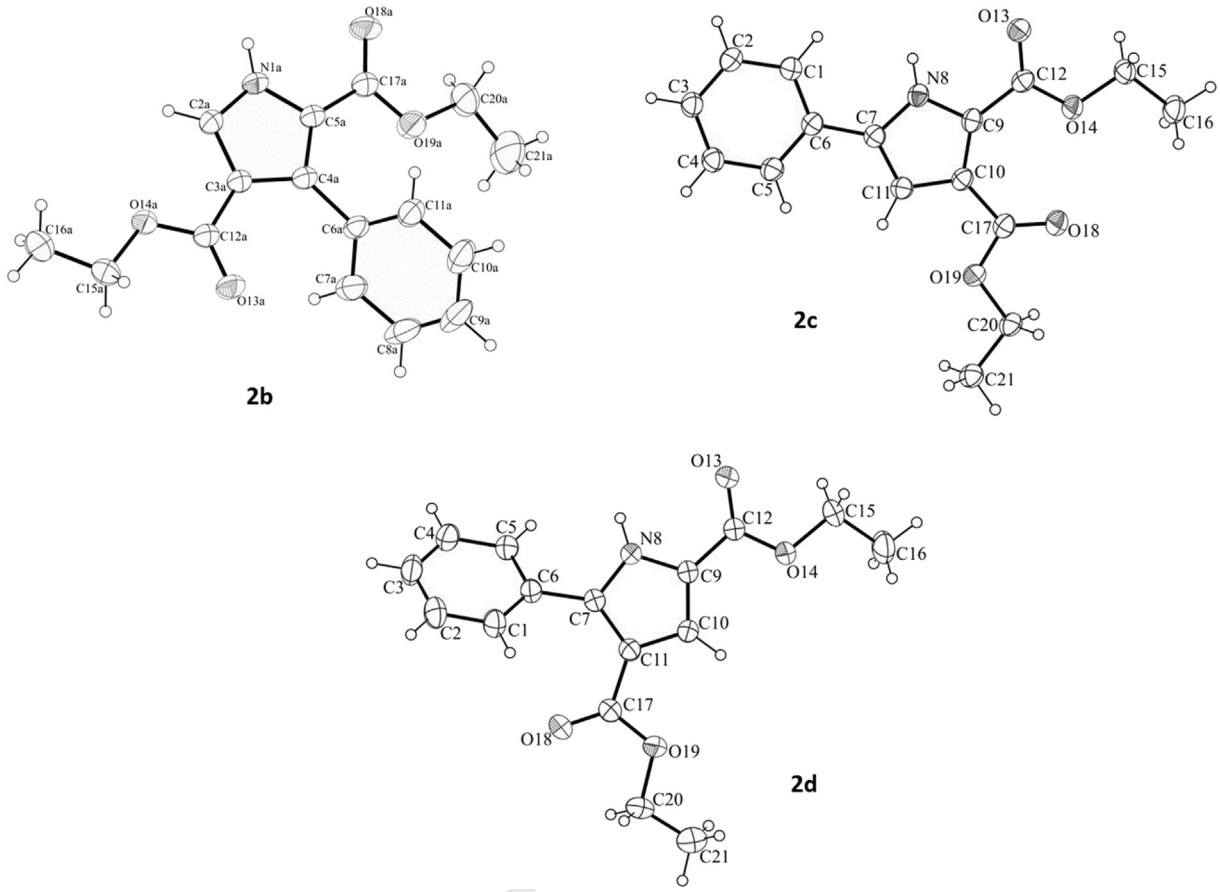


Fig. 2. The ORTEP drawing of phenyl-1H-pyrrole-diesters **2b-d** with thermal ellipsoids at 30% level.

most of the tested compounds showed antiproliferative activity with IC_{50} values from 3 to 24 μM, excepted **1a** and **1j** that were found inactive ($IC_{50} > 50$ μM). The two pyrroloquinoxalines **1d** and **1f** having an ester function and a phenyl respectively in position 2 and 3 exhibited better activities than their other homologues. Compounds **1d** and **1k** showed a better activity that the one noticed for the reference compound A6730: i.e. $IC_{50} = 3.5$ and 3 μM for **1d** and **1k**, respectively in comparison with 5.5 μM for A6730.

The antiproliferative activities of compounds **1a-l** against the T-acute lymphoblastic leukemia Jurkat cell line were also investigated, and the results exhibited potent cytotoxicity for pyrroloquinoxalines **1b-l** (IC_{50} from 2 to 6.5 μM) as potent as the one observed for A6730 ($IC_{50} = 3.5$ μM). Nevertheless, the pyrrolo[1,2-*a*]quinoxaline **1a** showed low antiproliferative activity ($IC_{50} = 41$ μM). Hence in the subseries functionalized by a phenyl and an ester at C2 and C3 respectively, the IC_{50} of **1b** and **1c** (5 μM) was 8.2 times lower than those of compound **1a** ($IC_{50} = 41$ μM).

Against the human myeloma cell line U266, the same pyrrolo[1,2-*a*]quinoxalines **1b**, **1f**, **1h**, **1k** and **1l**, bearing a benzylpiperidinyl fluorobenzimidazole or a benzylpiperidinyl triazolylpyridine moiety in position 4 and substituted on the pyrrole ring, exhibited potent cytotoxicity (IC_{50} from 3 to 5 μM). All the other tested pyrroloquinoxalines **1** (compounds **1d-e**, **1g** and **1i**) were also found to be active on U266 cell line with IC_{50} ranging from 8 to 18 μM, with the exception of **1a**, **1c** and **1j** that presented an IC_{50} superior to 50 μM.

Against each human cancer cell lines, the antiproliferative activities of compounds **1a-l** were generally found superior to those of the other reference drug LY-294002.

3.2. Cytotoxicity activity in activated normal peripheral blood mononuclear cells

The compounds **1a-l** were tested on activated (PBMNC + PHA) human peripheral blood mononuclear cells to evaluate their respective cytotoxicity on normal cells (Table 2). As expected, most of the pyrrolo[1,2-*a*]quinoxalines **1a-l** showed significant level of cytotoxicity against lymphocytes with IC_{50} ranging from 8 to over 50 μM. These preliminary results were used to determine their respective range of toxic concentration.

Indexes of selectivity (IS) were defined as the ratio of the IC_{50} value on the human mononuclear cells to the IC_{50} value on the K562, U937, HL60, Jurkat and U266 lines. Compounds that demonstrated high selectivity (high index of selectivity) should offer a potential of safer therapy. This led to identify compounds with index of selectivity >16.7 and >12.5 for compounds **1e** and **1a**, respectively, on the human myeloid leukemic cell lines K562; and >12.5 for compound **1g** against the human leukemic cell lines Jurkat. We could notice that the more interesting new pyrroloquinoxaline structure (compound **1e**) could be considered as a direct combination of our previously bioactive described derivatives JG572 and JG576. Moreover, we could notice that the compound **1a** showed interesting selectivity towards K562 CML cell lines. The potential inhibitor **1h** also showed interesting IS on U937 and U266 leukemic cell lines with value of 12.5. These four compounds could now constitute suitable candidates for further pharmacological studies. The reference compound A6730 showed interesting selectivity with index of selectivity value noticed at 14.3 on the Jurkat cell line.

Please cite this article in press as: V. Desplat, et al., Synthesis and evaluation of the cytotoxic activity of novel ethyl 4-[4-(4-substitutedpiperidin-1-yl)]benzyl-phenylpyrrolo[1,2-*a*]quinoxaline-carboxylate derivatives in myeloid and lymphoid leukemia cell lines, European Journal of Medicinal Chemistry (2016), http://dx.doi.org/10.1016/j.ejmech.2016.02.047

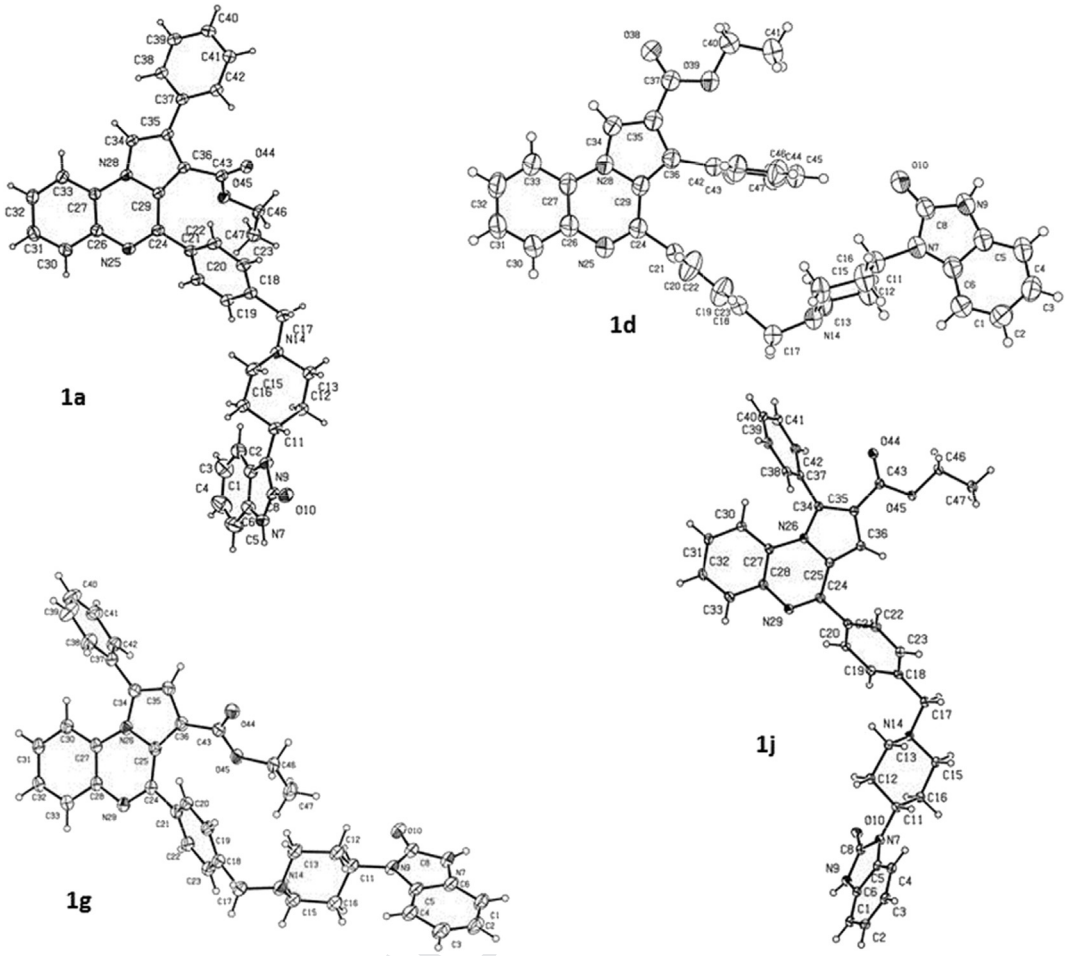


Fig. 3. The ORTEP drawing of pyrrolo[1,2-a]quinoxalines **1a, 1d, 1g** and **1j** with thermal ellipsoids at 30% level.

Table 2
In vitro activity of compounds **1a-l** on U937, K562, HL60, Jurkat and U266 cells, and cytotoxicity on human peripheral blood mononuclear cells PBMNC + PHA.

IC_{50} values (μM)[a]						
Compound	K562	U937	HL60	Jurkat	U266	Cytotoxicity on activated human peripheral blood mononuclear cells (PBMNC) PBMNC + PHA
A6730	17 ± 0.3	8 ± 0.2	5.5 ± 0.2	3.5 ± 0.2	n.d.[b]	>50
LY-294002	38 ± 1	14 ± 0.3	14 ± 0.3	22 ± 1	46 ± 2	>50
1a	4 ± 0.1	>50	>50	41 ± 1.2	>50	>50
1b	3 ± 0.1	21 ± 2.2	19 ± 2.8	5 ± 0.2	3.5 ± 0.1	14 ± 1
1c	14 ± 0.3	7 ± 0.3	12 ± 2.3	5 ± 0.1	>50	16 ± 2
1d	9 ± 0.3	>50	3.5 ± 0.1	4 ± 0.1	10 ± 0.8	41.3 ± 5
1e	3 ± 0.1	n.d.	10 ± 0.9	5 ± 0.2	18 ± 1.1	>50
1f	7 ± 0.2	7 ± 0.1	5 ± 0.1	2 ± 0.1	5 ± 0.1	8 ± 0.5
1g	8 ± 0.2	9 ± 0.4	6 ± 0.1	4 ± 0.1	9 ± 0.8	>50
1h	7 ± 0.3	4 ± 0.1	8 ± 0.2	6 ± 0.15	4 ± 0.1	50 ± 4
1i	12 ± 0.4	11 ± 0.9	24 ± 3	5 ± 0.1	8 ± 0.5	50 ± 6
1j	8.5 ± 0.3	>50	>50	6.5 ± 0.1	>50	>50
1k	3 ± 0.1	3 ± 0.1	3 ± 0.1	3.5 ± 0.1	3 ± 0.1	13 ± 1
1l	3.5 ± 0.1	8 ± 0.2	7 ± 0.3	3 ± 0.1	3 ± 0.1	12 ± 1

[a] The IC_{50} (μM) values correspond to the mean ± standard deviation from 3 independent experiments.
[b] n.d. = not determined.

3.3. Determination of the action mode

The reference compound A6730, which showed interesting selectivity on the Jurkat cell line, is additionally described as an Akt inhibitor (Akt1 IC_{50} = 58 nM, Akt2 IC_{50} = 10 nM and IC_{50} Akt3 = 2.2 μM) [31]. To determine the possible mechanism of action of our compounds, we first evaluated their potency on isolated enzymes such as Akt 1, 2 and 3 as well as on mTOR. These tests have been performed by DiscoverX at 1 and 10 μM [32]. Nevertheless, only low effects have been detected. Akt2 activity was reduced of 38% and 35% with derivatives **1e** and **1j** at 10 μM, respectively. Such of enzymatic interaction could not explain the cell effects. Next assessment of apoptosis in K562 cell line was examined. Apoptosis induces loss of membrane asymmetry resulting in phosphatidyl

serine (PS) exposure and alterations in mitochondrial membrane potential. Cells were incubated with or without increasing doses **1e** for 3 days. Compound **1e**, which showed the most interesting index of selectivity on the human myeloid leukemic cell lines K562, induces a significantly increase in Annexin V positive cells from the second and third days (92% ± 2.8) (Fig. 4). These results could explain the inhibition of cell proliferation observed with this compound **1e**.

3.4. Predicted toxicity and other drug relevant properties

To speculate on any toxicity risks and drug-like characteristics of these novel synthesized compounds, we computed a set of drug relevant properties from their 2D chemical structures (Table 3). The calculated properties indicated that our molecules **1** present the same toxicological profile as the anticancer reference compound A6730. These parameters were calculated by the molinspiration web services [33]. All compounds **1a-l** were found lipophilic with Clog *P* values between 6.44 and 7.70. The Clog *P* of the reference compound A6730 was noticed at 5.55 slightly lower than those of our compounds. To predict intestinal absorption, we estimated the molecular polar surface areas (PSA) of these new pyrroloquinoxaline compounds **1a-l** from the calculated TPSA. PSA has been extensively used in medicinal chemistry for modeling absorption phenomena and to optimize a drug's ability to permeate cells. The PSA of a molecule is defined as the surface sum over all polar atoms, which primarily consist of oxygen and nitrogen as well as their attached hydrogens. Molecules with a polar surface area greater than 140 $Å^2$ tend to be poor at permeating cell membranes, while a PSA of less than 60 $Å^2$ is usually needed for molecules to penetrate the blood-brain barrier and thus act on the brain and other central nervous system tissues [34,35]. The PSA of our compounds could suggest an intestinal absorption (PSA < 140 $Å^2$). On the other hand, these pyrroloquinoxalines seem to be unlikely to cross the blood-brain barrier (PSA > 60 $Å^2$). Thus, based on these predicted data, the calculated PSA values for these derivatives **1** could indicate a possible better oral bioavailability with a smaller chance of CNS toxicity.

4. Conclusion

In the present work, we synthesized a series of twelve new ethyl 4-[4-(4-substitutedpiperidin-1-yl)]benzyl-phenylpyrrolo[1,2-*a*] quinoxaline-carboxylate derivatives **1a-l** and investigated their antileukemic activity on the human leukemic cell lines U937, K562, Jurkat, U266 and HL60. These results have been discussed in a preliminary SAR study. The first biological evaluation of our new substituted pyrrolo[1,2-*a*]quinoxalines showed cytotoxic activity in these myeloid and lymphoid leukemia cell lines. Consequently, compounds **1e**, **1a**, **1g** and **1h** are promising due to their high cytotoxic activity against some leukemia cells (IC_{50} ranging from 3 to 9 μM) and their lower toxicity against normal hematopoietic cells (estimated IC_{50} > 50 μM). These compounds showing interesting index of selectivity may constitute suitable candidates for further pharmacological studies. Moreover, it would be also interesting to enlarge the biological evaluation of these new bioactive pyrrolo[1,2-*a*]quinoxaline derivatives in order to precise now their mechanism of action.

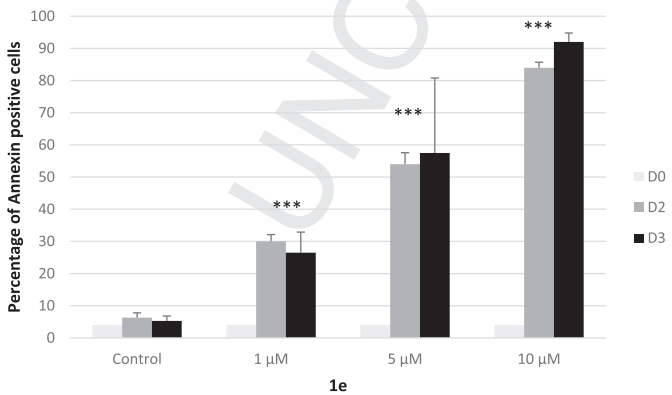


Fig. 4. Effect of **1e** on apoptosis of K562 cells. Cells were cultured with or without increasing doses of **1e** for 3 days in culture medium then stained with APC-Annexin V, and analyzed by flow cytometry. Results are expressed as Mean ± Standard Error (SEM) of three independent experiments. ***p < 0.001 compared with control (t-test).

Table 3
Predicted drug-relevant properties of compounds **1a-l**.

Compound	Clog *P*	TPSA	nON	nOH/NH	N violations[a]
A6730	5.55	95.50	8	2	2
1a	7.19	84.65	8	1	2
1b	7.70	75.53	7	1	2
1c	6.49	101.32	9	1	2
1d	7.19	84.65	8	1	2
1e	7.70	75.53	7	1	2
1f	6.49	101.32	9	1	2
1g	7.13	84.65	8	1	2
1h	7.65	75.53	7	1	2
1i	6.44	101.32	9	1	2
1j	7.13	84.65	8	1	2
1k	7.65	75.53	7	1	2
1l	6.44	101.32	9	1	2

[a] Number of violations to the Lipinski's "rule of five": log $P \leq 5$, molecular weight ≤500, number of hydrogen bond acceptors ≤10, and number of hydrogen bond donors ≤5.

5. Experimental

5.1. Chemistry

Commercially reagents were used as received without additional purification. Melting points were determined with an SM-LUX-POL Leitz hot-stage microscope and are uncorrected. IR spectra were recorded on an NICOLET 380FT-IR spectrophotometer. NMR spectra were recorded with tetramethylsilane as an internal standard using a BRUKER AVANCE 300 spectrometer. Splitting patterns have been designated as follows: s = singlet; bs = broad singlet; d = doublet; t = triplet; q = quartet; dd = double doublet; ddd = double double doublet; dt = double triplet; m = multiplet. Analytical TLC were carried out on 0.25 precoated silica gel plates (POLYGRAM SIL G/UV_{254}) and visualization of compounds after UV light irradiation. Silica gel 60 (70–230 mesh) was used for column chromatography. Microwave experiments were carried out using a focused microwave reactor (CEM Discover). High resolution mass spectra (electrospray in positive mode, ESI+) were recorded on a Waters Q-TOF Ultima apparatus. Elemental analyses were found within ±0.4% of the theoretical values.

5.1.1. Ethyl 2-azidoacrylate (**4**)

To a solution of sodium azide (3.75 g, 57.68 mmol) in DMF (120 mL) at 65 °C was added ethyl 2,3-dibromopropionate (5.00 g, 19.26 mmol). After 10 min, the reaction mixture was cooled and poured into water (300 mL) and extracted with ether (3 × 120 mL). The combined organic extracts were washed with water (3 × 120 mL), dried over $MgSO_4$, filtered and evaporated *in vacuo* to afford ethyl 2-azidoacrylate as a yellow oil (79%). R_f = 0.38 (cyclohexane/Et_2O-98/2) [36,37].

Please cite this article in press as: V. Desplat, et al., Synthesis and evaluation of the cytotoxic activity of novel ethyl 4-[4-(4-substitutedpiperidin-1-yl)]benzyl-phenylpyrrolo[1,2-*a*]quinoxaline-carboxylate derivatives in myeloid and lymphoid leukemia cell lines, European Journal of Medicinal Chemistry (2016), http://dx.doi.org/10.1016/j.ejmech.2016.02.047

*5.1.2. Diethyl 4-phenyl-1H-pyrrole-2,3-dicarboxylate (**2a**)*

To a 1,4-dioxane solution (10 mL) of 1,3-bis(diphenylphosphino) propane (dppp) (0.32 g, 0.78 mmol) were added ethyl isocyanoacetate (0.70 g, 6.19 mmol) and ethyl phenylpropiolate (0.90 g, 5.17 mmol). The solution was stirred at 100 °C for 2 h. After the consumption of ethyl isocyanoacetate, the reaction mixture was cooled to room temperature, filtered and evaporated *in vacuo*. The crude was purified on silica gel (eluent: cyclohexane/AcOEt-90/10 then cyclohexane/AcOEt-70/30, $R_f = 0.33$) to afford **2a** as a colorless oil (87%) [18,19].

*5.1.3. Diethyl 3-phenyl-1H-pyrrole-2,4-dicarboxylate (**2b**)*

Method A: To a mixture of ethyl phenylpropiolate (2.5 g, 14.3 mmol) and Ag_2CO_3 (0.264 g, 0.96 mmol) in 1,4-dioxane (40 mL) at 80 °C, ethyl isocyanoacetate (1.08 g, 9.55 mmol) was slowly added. The solution was stirred at 100 °C for 30 min. After the consumption of ethyl isocyanoacetate, the reaction mixture was cooled to room temperature, filtered and evaporated *in vacuo*. The crude was purified on silica gel (eluent: cyclohexane/AcOEt-90/10 then cyclohexane/AcOEt-70/30, $R_f = 0.33$) to afford **2b** as white crystals (9%). **Method B**: A mixture of ethyl phenylpropiolate (2.5 g, 14.3 mmol) and Ag_2CO_3 (0.264 g, 0.96 mmol) in 1,4-dioxane (40 mL) was heated at 80 °C for 5 min, then mixture was cooled with an ice bath. To this resulting solution cooled at room temperature, ethyl isocyanoacetate (1.08 g, 9.55 mmol) was added dropwise. The reaction mixture was stirred for 5 min at 25 °C, then heated for 30 min at 80 °C. The resulting slurry was concentrated under reduced pressure and taken up with dichloromethane. The organic layer was washed with brine, dried with Na_2SO_4, filtered and evaporated under *vacuum*. The crude was purified on silica gel as above leading to **2b** (10%). **Method C**: To a *N*-methyl-2-pyrrolidone (NMP) solution (40 mL) of Ag_2CO_3 (0.264 g, 0.96 mmol) was added ethyl phenylpropiolate (2.5 g, 14.3 mmol). After a prestirring of 30 s, the solution was irradiated during 1 min. The irradiation was programmed to maintain a constant temperature (80 °C) with a power of 200 W. To the resulting solution cooled at room temperature, ethyl isocyanoacetate (1.08 g, 9.55 mmol) was added dropwise. After a pre-stirring of 30 s, the solution was irradiated during 6 min with a first step of irradiation of 1 min at 25 °C followed by a second one of 5 min at 80 °C. The power was set at 200 W in both steps. The resulting slurry was concentrated under reduced pressure and taken up with dichloromethane. The organic layer was washed with brine, dried with Na_2SO_4, filtered and evaporated under *vacuum*. The crude was purified on silica gel as above to afford **2b** (12%). **Method D**: To a 1,4-dioxane solution (25 mL) of Cu_2O (0.085 g, 0.59 mmol) and 1,10-phenanthroline (0.215 g, 1.19 mmol) were added ethyl isocyanoacetate (1.35 g, 11.9 mmol) and ethyl phenylpropiolate (2.5 g, 14.3 mmol). After a prestirring of 30 s, the solution was irradiated during 40 min. The irradiation was programmed to maintain a constant temperature (100 °C) with a power of 200 W. The reaction mixture was cooled to room temperature, filtered and concentrated *in vacuo*. The crude was purified on silica gel as above to give **2b** (62%). **Method E**: To a THF solution (30 mL) of 1,8-diazabicyclo[5.4.0]undec-7-ene (DBU) (3.00 g, 19.7 mmol) was added ethyl isocyanoacetate (1.98 g, 17.5 mmol). After a prestirring of 30 s, the solution was irradiated during 2 min. The irradiation was programmed to maintain a constant temperature (50 °C) with a power of 200 W. To the resulting solution at the same temperature, ethyl isocyanoacetate (1.08 g, 9.55 mmol) was added dropwise. After a prestirring of 30 s, the solution was irradiated during 1 h. The irradiation was programmed to maintain a constant temperature (50 °C) with a power of 200 W. The reaction mixture was neutralized with acetic acid and then the solvent was removed under reduced pressure. The resulting residue was extracted with ethyl acetate and the extract was washed with hydrochloric acid and water, dried with Na_2SO_4, and then evaporated *in vacuo*. The crude was purified on silica gel as above to give **2b** (35%) [20–22].

*5.1.4. Diethyl 5-phenyl-1H-pyrrole-2,3-dicarboxylate (**2c**)*

To 1,4-diazabicyclo[2.2.2]octane (0.083 g, 0.74 mmol) and oxime **3** (1.00 g, 7.40 mmol) in dry toluene (20 mL) was added diethyl acetylenedicarboxylate (1.26 g, 7.40 mmol), and the resultant mixture was subjected to the two-stage microwave irradiation sequence (stage 1: 80 °C, 6 min; stage 2: 170 °C, 45 min). The power was set at 200 W in both steps. The reaction mixture was cooled to room temperature, filtered and concentrated *in vacuo*. The crude was purified on silica gel as above to give **2c** (28%) [24,25].

*5.1.5. Diethyl 5-phenyl-1H-pyrrole-2,4-dicarboxylate (**2d**)*

To a solution of ethyl 2-azidoacrylate **4** (1.70 g, 12.1 mmol) and ethyl phenylacetoacetate (3.47 g, 18.0 mmol) in MeOH (30 mL) was added AcOH (1.45 g, 24.1 mmol) and manganese(III) acetate dihydrate (1.29 g, 4.81 mmol), and the solution was stirred at 40 °C for 2 h. The reaction mixture was quenched with pH 9 ammonium buffer ($AcONH_4 + NH_4OH$), and then extracted twice with AcOEt. The combined organic extracts were washed with brine, dried with Na_2SO_4, filtered and concentrated *in vacuo*. The resulting residue was cooled, triturated in Et_2O and filtered. The crystals formed were filtered, and dried under reduced pressure to give **2d** as white crystals (42%) [26].

*5.1.6. General procedure for diethyl 1-(2-nitrophényl)-phenyl-pyrrole-dicarboxylate (**5a-d**)*

Conventional heating: To a solution of diethyl phenyl-1*H*-pyrrole-dicarboxylate **2a-d** (2.75 mmol) in 11 mL of DMF was added cesium carbonate (3.3 mmol). The mixture was stirred at room temperature for 10 min, then 2-fluoro-nitrobenzene (4.12 mmol) was added. The reaction mixture was refluxed for 1 h 30 (15 h for **5d**), then was diluted in AcOEt (35 mL). The organic layer was washed with water (2 × 30 mL), then brine (30 mL) and dried over sodium sulfate. The organic layer was concentrated under vacuo to give a brown oil. After triturating in Et_2O a solid was obtained and filtered off, washed with Et_2O and dried to give the desired product **5**. *Microwave heating:* A suspension of diethyl phenyl-1*H*-pyrrole-dicarboxylate **2a-d** (3.4 mmol), 1-fluoro-2-nitrobenzene (5.1 mmol) and cesium carbonate (4.06 mmol) in 12 mL of DMF was irradiated during 10 min. The irradiation was programmed to maintain a constant temperature (150 °C) with a maximal power output of 200 W. The reaction mixture was then diluted in AcOEt (60 mL), washed with water (2 × 50 mL), then brine (50 mL) and dried over sodium sulphate. The organic layer was concentrated under vacuo to give products **5a-d** as an oil.

*5.1.6.1. Ethyl 1-(2-nitrophenyl)-4-phenyl-pyrrole-2,3-dicarboxylate (**5a**).* Orange oil (85%/89%). ^{1}H NMR ($CDCl_3$) δ: 8.13 (dd, 1H, $J = 7.90$ and 1.35 Hz, H-3′), 7.74 (ddd, 1H, $J = 7.90$, 7.90 and 1.35 Hz, H-4′), 7.63 (ddd, 1H, $J = 7.90$, 7.90 and 1.35 Hz, H-5′), 7.55–7.43 (m, 3H, H-6′, H-2 phenyl and H-6 phenyl), 7.40–7.28 (m, 3H, H-3 phenyl, H-4 phenyl and H-5 phenyl), 7.02 (s, 1H, H-5), 4.39–4.30 (m, 2H, OCH_2), 4.12 (q, 2H, $J = 7.20$ Hz, OCH_2), 1.31 (t, 3H, $J = 7.20$ Hz, CH_3), 1.15 (t, 3H, $J = 7.20$ Hz, CH_3). HRMS-ESI *m/z* $[M + Na]^+$ Calcd for $C_{22}H_{20}N_2O_6Na$: 431.1219, Found: 431.1224.

*5.1.6.2. Ethyl 1-(2-nitrophenyl)-3-phenyl-pyrrole-2,4-dicarboxylate (**5b**).* Yellow oil (89%/95%). ^{1}H NMR ($CDCl_3$) δ: 8.19 (dd, 1H, $J = 7.80$ and 1.50 Hz, H-3′), 7.77 (ddd, 1H, $J = 7.80$, 7.80 and 1.50 Hz, H-4′), 7.67 (ddd, 1H, $J = 7.80$, 7.80 and 1.50 Hz, H-5′), 7.59 (s, 1H, H-5), 7.51 (dd, 1H, $J = 7.80$ and 1.50 Hz, H-6′), 7.39–7.35 (m, 5H, 5H phenyl), 4.14 (q, 2H, $J = 7.20$ Hz, OCH_2), 3.89–3.80 (m, 2H, OCH_2),

Please cite this article in press as: V. Desplat, et al., Synthesis and evaluation of the cytotoxic activity of novel ethyl 4-[4-(4-substitutedpiperidin-1-yl)]benzyl-phenylpyrrolo[1,2-*a*]quinoxaline-carboxylate derivatives in myeloid and lymphoid leukemia cell lines, European Journal of Medicinal Chemistry (2016), http://dx.doi.org/10.1016/j.ejmech.2016.02.047

1.13 (t, 3H, $J = 7.20$ Hz, CH_3), 0.77 (t, 3H, $J = 7.20$ Hz, CH_3). HRMS-ESI m/z $[M + Na]^+$ Calcd for $C_{22}H_{20}N_2O_6Na$: 431.1219, Found: 431.1236.

*5.1.6.3. Ethyl 1-(2-nitrophenyl)-5-phenyl-pyrrole-2,3-dicarboxylate (**5c**).* Yellow oil (65%/76%). ^1H NMR ($CDCl_3$) δ: 8.05–8.01 (m, 2H, H-3′ and H-6′), 7.62–7.56 (m, 2H, H-2 phenyl and H-6 phenyl), 7.31 (t, 1H, $J = 7.50$ Hz, H-4′), 7.25–7.04 (m, 4H, H-5′, H-3 phenyl, H-4 phenyl and H-5 phenyl), 6.77 (s, 1H, H-4), 4.34 (q, 2H, $J = 6.90$ Hz, OCH_2), 4.12 (q, 2H, $J = 6.90$ Hz, OCH_2), 1.38–1.14 (m, 6H, $2CH_3$). HRMS-ESI m/z $[M + Na]^+$ Calcd for $C_{22}H_{20}N_2O_6Na$: 431.1219, Found: 431.1248.

*5.1.6.4. Ethyl 1-(2-nitrophenyl)-5-phenyl-pyrrole-2,4-dicarboxylate (**5d**).* Yellow oil (28%/). ^1H NMR ($CDCl_3$) δ: 8.04 (dd, 1H, $J = 7.50$ and 1.80 Hz, H-3′), 7.66 (s, 1H, H-3), 7.52 (ddd, 1H, $J = 7.50$, 7.50 and 1.80 Hz, H-4′), 7.46 (ddd, 1H, $J = 7.50$, 7.50 and 1.80 Hz, H-5′), 7.26–7.16 (m, 6H, H-6′ and 5H phenyl), 4.17 (q, 2H, $J = 6.90$ Hz, OCH_2), 4.15 (q, 2H, $J = 6.90$ Hz, OCH_2), 1.27 (t, 3H, $J = 6.90$ Hz, CH_3), 1.18 (t, 3H, $J = 6.90$ Hz, CH_3). HRMS-ESI m/z $[M+Na]^+$ Calcd for $C_{22}H_{20}N_2O_6Na$: 431.1219, Found: 431.1240.

*5.1.7. General procedure for ethyl 4,5-dihydro-4-oxo-5H-phénylpyrrolo[1,2-a]quinoxaline-carboxylate (**6a-d**)*

A suspension of **5** (2.6 mmol) and iron powder (9.5 mmol) in 11 mL of acetic acid was heated under reflux for 2 h. The reaction mixture was cooled, suspended in 20 mL of a 1 M aqueous solution of HCl, agitated, then filtered off, washed with HCl 1 M (8 mL), water, Et_2O and dried to give **6** as a fluffy white solid.

*5.1.7.1. Ethyl 4,5-dihydro-4-oxo-5H-2-phenylpyrrolo[1,2-a]quinoxaline-3-carboxylate (**6a**).* Beige crystals (69%), mp 239–241 °C. IR (KBr) 3200, 2750 (NH), 1720 (COO), 1660 (CON). ^1H NMR (DMSO-d_6) δ: 11.49 (s, 1H, NH), 8.58 (s, 1H, H-1), 8.17 (d, 1H, $J = 8.10$ Hz, H-9), 7.55 (d, 2H, $J = 7.50$ Hz, H-2 phenyl and H-6 phenyl), 7.45 (t, 2H, $J = 7.50$ Hz, H-3 phenyl and H-5 phenyl), 7.34–7.27 (m, 4H, H-4 phenyl, H-6, H-7 and H-8), 4.30 (q, 2H, $J = 7.20$ Hz, OCH_2), 1.25 (t, 3H, $J = 7.20$ Hz, CH_3). ^{13}C NMR (DMSO-d_6) δ: 157.0 (C-4), 155.2 (C-7), 129.7 (C-9a), 122.6 (C-5a), 117.7 (C-1), 116.8 (C-3a), 116.0 (C-6), 112.3 (C-3), 111.0 (C-2), 108.9 (C-9), 100.8 (C-8), 55.4 (CH_3O). Anal. Calcd. for $C_{12}H_{10}N_2O_2$: C, 67.28; H, 4.71; N, 13.08. Found: C, 67.40; H, 4.81; N, 12.94.

*5.1.7.2. Ethyl 4,5-dihydro-4-oxo-5H-3-phenylpyrrolo[1,2-a]quinoxaline-2-carboxylate (**6b**).* Beige crystals (88%), mp > 310 °C. IR (KBr) 3200, 2800 (NH), 1720 (COO), 1665 (CON). ^1H NMR (DMSO-d_6) δ: 11.25 (s, 1H, NH), 8.80 (s, 1H, H-1), 8.28 (d, 1H, $J = 7.80$ Hz, H-9), 7.34–7.27 (m, 7H, H-6, H-8 and 5H phenyl), 7.21 (t, 1H, $J = 7.80$ Hz, H-7), 4.09 (q, 2H, $J = 6.90$ Hz, OCH_2), 1.09 (t, 3H, $J = 6.90$ Hz, CH_3). ^{13}C NMR (DMSO-d_6) δ: 157.0 (C-4), 155.2 (C-7), 129.7 (C-9a), 122.6 (C-5a), 117.7 (C-1), 116.8 (C-3a), 116.0 (C-6), 112.3 (C-3), 111.0 (C-2), 108.9 (C-9), 100.8 (C-8), 55.4 (CH_3O). Anal. Calcd. for $C_{12}H_{10}N_2O_2$: C, 67.35; H, 4.78; N, 13.15. Found: C, 67.47; H, 4.92; N, 13.30.

*5.1.7.3. Ethyl 4,5-dihydro-4-oxo-5H-1-phenylpyrrolo[1,2-a]quinoxaline-3-carboxylate (**6c**).* White crystals (58%), mp = 218–220 °C. IR (KBr) 3200, 2700 (NH), 1700 (COO), 1670 (CON). ^1H NMR (DMSO-d_6) δ: 11.50 (s, 1H, NH), 7.56–7.53 (m, 5H, 5H phenyl), 7.32 (d, 1H, $J = 7.70$ Hz, H-9), 7.24 (t, 1H, $J = 7.70$ Hz, H-8), 6.97 (d, 1H, $J = 7.70$ Hz, H-6), 6.84 (t, 1H, $J = 7.70$ Hz, H-7), 6.81 (s, 1H, H-2), 4.27 (q, 2H, $J = 6.90$ Hz, OCH_2), 1.30 (t, 3H, $J = 6.90$ Hz, CH_3). Anal. Calcd. for $C_{12}H_{10}N_2O_2$: C, 67.21; H, 4.64; N, 13.01. Found: C, 67.33; H, 4.74; N, 12.76.

*5.1.7.4. Ethyl 4,5-dihydro-4-oxo-5H-1-phenylpyrrolo[1,2-a]quinoxaline-2-carboxylate (**6d**).* White crystals (84%), mp = 229–231 °C. IR (KBr) 3200, 2700 (NH), 1695 (COO), 1665 (CON). ^1H NMR (DMSO-d_6) δ: 11.53 (s, 1H, NH), 7.61–7.53 (m, 3H, 3H phenyl), 7.50–7.47 (m, 2H, 2H phenyl), 7.44 (s, 1H, H-3), 7.29 (d, 1H, $J = 7.90$ Hz, H-6), 7.20 (t, 1H, $J = 7.90$ Hz, H-8), 6.75 (t, 1H, $J = 7.90$ Hz, H-7), 6.64 (d, 1H, $J = 7.90$ Hz, H-9), 4.02 (q, 2H, $J = 6.90$ Hz, OCH_2), 1.30 (t, 3H, $J = 6.90$ Hz, CH_3). Anal. Calcd. for $C_{12}H_{10}N_2O_2$: C, 67.31; H, 4.74; N, 13.11. Found: C, 67.43; H, 4.95; N, 12.86.

*5.1.8. General procedure for ethyl 4-chloro-phenylpyrrolo[1,2-a] quinoxaline-carboxylate (**7a-d**)*

A solution of 5*H*-pyrrolo[1,2-a]quinoxalin-4-one **6** (4 mmol) in $POCl_3$ (8 mL) was refluxed for 2 h. After removing excess of reactive under vacuum, the residue was carefully dissolved in water at 0 °C and the resulting solution was made basic with sodium carbonate. The precipitate was filtered, dried and recrystallized from ethyl acetate to give **7**.

*5.1.8.1. Ethyl 4-chloro-2-phenylpyrrolo[1,2-a]quinoxaline-3-carboxylate (**7a**).* White crystals (77%), mp = 109–111 °C. IR (KBr) 1715 (COO). ^1H NMR ($CDCl_3$) δ: 7.97 (s, 1H, H-1), 7.87 (dd, 1H, $J = 8.10$ and 1.20 Hz, H-9), 7.80 (dd, 1H, $J = 8.10$ and 1.20 Hz, H-6), 7.58–7.32 (m, 7H, H-7, H-8 and 5H phenyl), 4.41 (q, 2H, $J = 7.20$ Hz, OCH_2), 1.34 (t, 3H, $J = 7.20$ Hz, CH_3). HRMS-ESI m/z $[M+H]^+$ Calcd for $C_{20}H_{15}N_2O_2ClNa$: 351.0900, Found: 351.0918.

*5.1.8.2. Ethyl 4-chloro-3-phenylpyrrolo[1,2-a]quinoxaline-2-carboxylate (**7b**).* Beige crystals (96%), mp = 144–146 °C. IR (KBr) 1715 (COO). ^1H NMR ($CDCl_3$) δ: 8.56 (s, 1H, H-1), 7.93 (dd, 1H, $J = 8.10$ and 1.50 Hz, H-9), 7.88 (dd, 1H, $J = 8.10$ and 1.50 Hz, H-6), 7.59 (ddd, 1H, $J = 8.10$, 7.90 and 1.50 Hz, H-8), 7.52 (ddd, 1H, $J = 8.10$, 7.90 and 1.50 Hz, H-7), 7.44–7.37 (m, 5H, 5H phenyl), 4.19 (q, 2H, $J = 7.20$ Hz, OCH_2), 1.14 (t, 3H, $J = 7.20$ Hz, CH_3). HRMS-ESI m/z $[M+H]^+$ Calcd for $C_{20}H_{15}N_2O_2ClNa$: 351.0900, Found: 351.0933.

*5.1.8.3. Ethyl 4-chloro-1-phenylpyrrolo[1,2-a]quinoxaline-3-carboxylate (**7c**).* Beige crystals (94%), mp = 108–110 °C. IR (KBr) 1720 (COO). ^1H NMR ($CDCl_3$) δ: 7.91 (d, 1H, $J = 8.10$ Hz, H-9), 7.55–7.48 (m, 5H, 5H phenyl), 7.41 (t, 1H, $J = 8.10$ Hz, H-8), 7.33 (d, 1H, $J = 8.10$ Hz, H-6), 7.17 (t, 1H, $J = 8.10$ Hz, H-7), 7.09 (s, 1H, H-2), 4.44 (q, 2H, $J = 7.20$ Hz, OCH_2), 1.44 (t, 3H, $J = 7.20$ Hz, CH_3). HRMS-ESI m/z $[M+H]^+$ Calcd for $C_{20}H_{15}N_2O_2ClNa$: 351.0900, Found: 351.0903.

*5.1.8.4. Ethyl 4-chloro-1-phenylpyrrolo[1,2-a]quinoxaline-2-carboxylate (**7d**).* Orange crystals (90%), mp = 178 °C. IR (KBr) 1710 (COO). ^1H NMR ($CDCl_3$) δ: 7.86 (dd, 1H, $J = 8.10$ and 1.30 Hz, H-6), 7.63–7.55 (m, 4H, 3H phenyl and H-3), 7.50–7.46 (m, 2H, 2H phenyl), 7.37 (ddd, 1H, $J = 8.10$, 7.85 and 1.30 Hz, H-8), 7.10 (ddd, 1H, $J = 8.10$, 7.85 and 1.30 Hz, H-7), 7.01 (dd, 1H, $J = 8.10$ and 1.30 Hz, H-9), 4.19 (q, 2H, $J = 7.20$ Hz, OCH_2), 1.18 (t, 3H, $J = 7.20$ Hz, CH_3). HRMS-ESI m/z $[M + H]^+$ Calcd for $C_{20}H_{15}N_2O_2ClNa$: 351.0900, Found: 351.0925.

*5.1.9. General procedure for ethyl 4-(4-formylphenyl)-phenylpyrrolo[1,2-a]quinoxaline-carboxylate (**8a-d**)*

To suspension of compound **7a-d** (4.64 mmol), and $Pd(PPh_3)_4$ (0.232 mmol) in a mixture of toluene/EtOH (75/4.1 mL) under nitrogen were added K_2CO_3 (5.1 mmol) and 4-formylphenylboronic acid (5.1 mmol). The reaction mixture was refluxed for 24 h, and the cooled suspension was extracted with CH_2Cl_2 (3 × 80 mL). The organic layer was washed with a saturated solution of NaCl (95 mL), and the combined organic extracts were dried over sodium sulfate, filtered, and evaporated under reduced pressure. The crude residue was triturated in ethanol. The resulting precipitate was filtered, washed with ethanol, and purified by column chromatography on

Please cite this article in press as: V. Desplat, et al., Synthesis and evaluation of the cytotoxic activity of novel ethyl 4-[4-(4-substitutedpiperidin-1-yl)]benzyl-phenylpyrrolo[1,2-*a*]quinoxaline-carboxylate derivatives in myeloid and lymphoid leukemia cell lines, European Journal of Medicinal Chemistry (2016), http://dx.doi.org/10.1016/j.ejmech.2016.02.047

silica gel using dichloromethane as eluent gave the pure product **8**.

*5.1.9.1. Ethyl 4-(4-formylphenyl)-2-phenylpyrrolo[1,2-a]quinoxaline-3-carboxylate (**8a**).* Yellow crystals (83%), mp = 76–78 °C. IR (KBr) 1720 (COO), 1700 (CHO). ^{1}H NMR ($CDCl_3$) δ: 10.13 (s, 1H, CHO), 8.12 (dd, 1H, *J* = 8.10 and 1.30 Hz, H-9), 8.10 (s, 1H, H-1), 8.04 (d, 2H, *J* = 7.80 Hz, H-3′ and H-5′), 8.00 (dd, 1H, *J* = 8.10 and 1.30 Hz, H-6), 7.96 (d, 2H, *J* = 7.80 Hz, H-2′ and H-6′), 7.67 (ddd, 1H, *J* = 8.10, 7.85 and 1.30 Hz, H-8), 7.64–7.52 (m, 3H, H-7 and 2H phenyl), 7.48–7.39 (m, 3H, 3H phenyl), 3.63 (q, 2H, *J* = 6.90 Hz, OCH_2), 0.89 (t, 3H, *J* = 6.90 Hz, CH_3). HRMS-ESI *m/z* $[M + H]^+$ Calcd for $C_{27}H_{20}N_2O_3Na$: 421.1552, Found: 421.1552.

*5.1.9.2. Ethyl 4-(4-formylphenyl)-3-phenylpyrrolo[1,2-a]quinoxaline-2-carboxylate (**8b**).* Yellow crystals (60%), mp = 160–162 °C. IR (KBr) 1715 (COO), 1700 (CHO). ^{1}H NMR ($CDCl_3$) δ: 9.93 (s, 1H, CHO), 8.69 (s, 1H, H-1), 8.09 (d, 1H, *J* = 7.80 Hz, H-9), 8.05 (d, 1H, *J* = 7.80 Hz, H-6), 7.66 (t, 1H, *J* = 7.80 Hz, H-8), 7.58 (t, 1H, *J* = 7.80 Hz, H-7), 7.53 (d, 2H, *J* = 7.20 Hz, H-3′ and H-5′), 7.38 (d, 2H, *J* = 7.20 Hz, H-2′ and H-6′), 7.03–6.95 (m, 5H, 5H phenyl), 4.23 (q, 2H, *J* = 6.90 Hz, OCH_2), 1.20 (t, 3H, *J* = 6.90 Hz, CH_3). HRMS-ESI *m/z* $[M+H]^+$ Calcd for $C_{27}H_{20}N_2O_3Na$: 421.1552, Found: 421.1532.

*5.1.9.3. Ethyl 4-(4-formylphenyl)-1-phenylpyrrolo[1,2-a]quinoxaline-3-carboxylate (**8c**).* Pale yellow crystals (79%), mp = 178–180 °C. IR (KBr) 1720 (COO), 1695 (CHO). ^{1}H NMR ($CDCl_3$) δ: 10.14 (s, 1H, CHO), 8.09–8.06 (m, 1H, H-9), 8.04 (d, 2H, *J* = 8.10 Hz, H-3′ and H-5′), 7.97 (d, 2H, *J* = 8.10 Hz, H-2′ and H-6′), 7.58–7.56 (m, 5H, 5H phenyl), 7.48–7.44 (m, 2H, H-6 and H-8), 7.25–7.20 (m, 1H, H-7), 7.21 (s, 1H, H-2), 3.80 (q, 2H, *J* = 7.20 Hz, OCH_2), 0.99 (t, 3H, *J* = 7.20 Hz, CH_3). HRMS-ESI *m/z* $[M + H]^+$ Calcd for $C_{27}H_{20}N_2O_3Na$: 421.1552, Found: 421.1573.

*5.1.9.4. Ethyl 4-(4-formylphenyl)-1-phenylpyrrolo[1,2-a]quinoxaline-2-carboxylate (**8d**).* Yellow crystals (89%), mp = 182–184 °C. IR (KBr) 1700 (COO and CHO). ^{1}H NMR ($CDCl_3$) δ: 10.18 (s, 1H, CHO), 8.20 (d, 2H, *J* = 7.95 Hz, H-3′ and H-5′), 8.12 (d, 2H, *J* = 7.95 Hz, H-2′ and H-6′), 8.02 (d, 1H, *J* = 7.80 Hz, H-6), 7.62–7.48 (m, 6H, 5H phenyl and H-3), 7.42 (t, 1H, *J* = 7.80 Hz, H-8), 7.13–7.09 (m, 2H, H-7 and H-9), 4.16 (q, 2H, *J* = 7.20 Hz, OCH_2), 1.13 (t, 3H, *J* = 7.20 Hz, CH_3). HRMS-ESI *m/z* $[M + H]^+$ Calcd for $C_{27}H_{20}N_2O_3Na$: 421.1552, Found: 421.1566.

*5.1.10. General procedure for ethyl 4-{4-[(4-(2-oxo-2,3-dihydro-1H-benzimidazol-1-yl)piperidin-1-yl)benzyl]}-phenylpyrrolo[1,2-a]quinoxaline-carboxylate, ethyl 4-{4-[(4-(5-fluoro-1H-benzimidazol-2-yl)piperidin-1-yl)benzyl]}-phenylpyrrolo[1,2-a]quinoxaline-carboxylate and ethyl 4-{4-[4-(3-(pyridin-2-yl)-1,2,4-triazol-5-yl)piperidin-1-yl)benzyl]}-phenylpyrrolo[1,2-a]quinoxaline-carboxylate (**1a-l**)*

The pH of a solution of the aldehyde **8a-d** (0.784 mmol) and 4-(2-ketobenzimidazolin-1-yl)piperidine or 4-(5-chloro-2-ketobenzimidazolin-1-yl)piperidine or 2-(3-piperidin-4-yl-1*H*-1,2,4-triazol-5-yl)pyridine (0.941 mmol) in 15 mL methanol was adjusted to 6 by the dropwise addition of acetic acid. Powered sodium cyanoborohydride (2.15 mmol) was then added, and the resultant mixture was refluxed for 5 h. After removal of the methanol by rotary evaporation, the residue was triturated in water and extracted with dichloromethane. The organic layer was washed with water, dried over magnesium sulfate and evaporated to dryness. Column chromatography of the residue on silica gel using ethyl acetate - cyclohexane (1/1) then methanol-chloroform (1/9) as eluents gave the crude product. This solid was then triturated with diethyl ether, filtered, washed with diethyl ether and dried under reduced pressure to give the compounds **1a-l**.

*5.1.10.1. Ethyl 4-{4-[(4-(2-oxo-2,3-dihydro-1H-benzimidazol-1-yl)piperidin-1-yl)benzyl]}-2-phenylpyrrolo[1,2-a]quinoxaline-3-carboxylate (**1a**).* White crystals (53%), mp 185–187 °C. ^{1}H NMR ($CDCl_3$) δ: 9.68 (s, 1H, NH), 8.10 (d, J = 7.8 Hz, 1H, H-9), 8.08 (s, 1H, H-1), 7.96 (d, J = 7.5 Hz, 1H, H-6), 7.78 (d, J = 7.8 Hz, 2H, H-2′ and H-6′), 7.64–7.53 (m, 6H, H-7, H-8, H-3′, H-5′, H-4 benzimid. and H-7 benzimid.), 7.45–7.30 (m, 4H, H-2 phenyl, H-3 phenyl, H-5 phenyl and H-6 phenyl), 7.10–7.07 (m, 3H, H-4 phenyl, H-5 benzimid. and H-6 benzimid.), 4.46–4.37 (m, 1H, CH pip.), 3.67 (q, J = 6.9 Hz, 2H, CH_2), 3.66 (s, 2H, CH_2N), 3.13–3.08 (m, 2H, NCH_2 pip.), 2.58–2.44 (m, 2H, NCH_2 pip.), 2.28–2.20 (m, 2H, CH_2 pip.), 1.88–1.84 (m, 2H, CH_2 pip.), 0.92 (t, J = 6.9 Hz, 3H, CH_3). ^{13}C NMR ($CDCl_3$) δ: 166.94 (C═O), 156.52 (C-4), 155.69 (C═O benzimid.), 141.51 (C-3a and C-5a), 139.74 (C-4′ and C-1 phenyl), 137.42 (C-7a benzimid.), 134.64 (C-1′), 131.82 (C-6), 130.59 (C-3 and C-5 phenyl), 130.37 (C-3′ and C-5′), 129.86 (C-2′ and C-6′), 129.71 (C-4 phenyl), 129.33 (C-2 and C-6 phenyl), 129.04 (C-7), 127.57 (C-8), 127.19 (C-2 and C-3a benzimid.), 125.08 (C-9a), 122.49 (C-5 benzimid. and C-6 benzimid.), 115.18 (C-9), 114.61 (C-3), 114.37 (C-1), 111.15 (C-4 benzimid. and C-7 benzimid.), 64.05 (NCH_2), 62.47 (CH_2), 54.56 (NCH_2 pip.), 52.12 (CH pip.), 30.65 (CH_2 pip.), 15.08 (CH_3). HRMS-ESI *m/z* $[M+Na]^+$ Calcd for $C_{39}H_{35}N_5O_3Na$: 644.2637, Found: 644.2634.

*5.1.10.2. Ethyl 4-{4-[(4-(5-fluoro-1H-benzimidazol-2-yl)piperidin-1-yl)benzyl]}-2-phenylpyrrolo[1,2-a]quinoxaline-3-carboxylate (**1b**).* Cream crystals (61%), mp 162–164 °C. ^{1}H NMR ($CDCl_3$) δ: 8.08 (s, 1H, H-1), 8.04 (d, J = 7.8 Hz, 1H, H-9), 7.96 (d, J = 7.8 Hz, 1H, H-6), 7.70 (d, J = 7.2 Hz, 2H, H-2′ and H-6′), 7.59 (t, J = 7.2 Hz, 1H, H-7), 7.55–7.47 (m, 3H, H-3′, H-5′ and H-8), 7.44–7.33 (m, 7H, H-2 phenyl, H-3 phenyl, H-5 phenyl, H-6 phenyl, H-4 benzimid., H-6 benzimid. and H-7 benzimid.), 6.93 (t, J = 7.2 Hz, 1H, H-4 phenyl), 3.61 (q, J = 7.0 Hz, 2H, CH_2), 3.54 (s, 2H, CH_2N), 2.91 (d, 2H, J = 10.2 Hz, NCH_2 pip.), 2.84 (t, J = 11.4 Hz, 1H, CH pip.), 2.05 (t, J = 10.8 Hz, 2H, CH_2 pip.), 1.99 (d, J = 11.4 Hz, 2H, NCH_2 pip.), 1.94–1.85 (m, 2H, CH_2 pip.), 0.87 (t, J = 7.0 Hz, 3H, CH_3). ^{13}C NMR ($CDCl_3$) δ: 165.68 (C═O), 160.05 (d, J = 234 Hz, C-5 benzimid.), 158.46 (C-4), 139.78 (C-4′), 138.14 (C-1 phenyl, C-1′, C-2 benzimid., C-5a and C-3a benzimid.), 135.91 (C-3a and C-5a), 133.21 (C-6 and C-7a benzimid.), 130.55 (C-2′ and C-6′), 129.99 (C-3′ and C-5′), 129.23 (C-7), 128.91 (C-8), 128.51 (C-4 phenyl), 127.95 (C-2 phenyl, C-3 phenyl, C-5 phenyl and C-6 phenyl), 127.73 (C-9a), 125.85 (C-2), 123.62 (C-3), 113.95 (C-9), 113.48 (C-1 and C-7 benzimid.), 113.24 (C-4 benzimid. and C-6 benzimid.), 62.71 (NCH_2), 61.16 (CH_2), 53.17 (NCH_2 pip.), 36.50 (CH pip.), 30.74 (CH_2 pip.), 13.63 (CH_3). HRMS-ESI *m/z* $[M+H]^+$ Calcd for $C_{39}H_{35}N_5O_2F$: 624.2775, Found: 624.2751.

*5.1.10.3. Ethyl 4-{4-[4-(3-(pyridin-2-yl)-1,2,4-triazol-5-yl)piperidin-1-yl)benzyl]}-2-phenylpyrrolo[1,2-a]quinoxaline-3-carboxylate (**1c**).* White crystals (56%), mp 142–144 °C. ^{1}H NMR ($CDCl_3$) δ: 8.73 (d, J = 3.0 Hz, 1H, H-6 pyr.), 8.22 (d, J = 7.8 Hz, 1H, H-5 pyr.), 8.08 (d, J = 8.4 Hz, 1H, H-6), 8.06 (s, 1H, H-1), 7.94 (d, J = 8.4 Hz, 1H, H-9), 7.84 (t, J = 7.8 Hz, 1H, H-4 pyr.), 7.74 (d, J = 7.8 Hz, 2H, H-2′ and H-6′), 7.58 (t, J = 7.8 Hz, 1H, H-7), 7.56–7.49 (m, 4H, H-3′, H-5′, H-2 phenyl and H-6 phenyl), 7.44–7.40 (m, 2H, H-3 phenyl and H-5 phenyl), 7.38–7.34 (m, 2H, H-4 phenyl and H-3 pyr.), 7.28 (m, 1H, H-8), 3.65 (s, 2H, CH_2N), 3.64 (q, J = 7.2 Hz, 2H, CH_2), 3.06 (d, 2H, J = 10.8 Hz, NCH_2 pip.), 2.97–2.89 (m, 1H, CH pip.), 2.23 (t, J = 10.2 Hz, 2H, CH_2 pip.), 2.12 (d, J = 12.0 Hz, 2H, NCH_2 pip.), 2.10–2.00 (m, 2H, CH_2 pip.), 0.91 (t, J = 7.2 Hz, 3H, CH_3). ^{13}C NMR ($CDCl_3$) δ: 165.56 (C═O), 154.36 (C-4, C-3 Triazole and C-2 pyr.), 149.50 (C-6 pyr.), 147.15 (C-5 Triazole), 138.36 (C-4′ and C-4 pyr.), 137.48 (C-1 phenyl, C-1′ and C-5a), 136.04 (C-3a), 133.31 (C-6), 130.40 (C-2′ and C-6′), 129.41 (C-3′, C-5′ and C-7), 128.97 (C-8), 128.47 (C-4 phenyl), 128.31 (C-2 phenyl and C-6 phenyl), 127.95 (C-3 phenyl and C-5 phenyl), 127.62 (C-9a), 125.83 (C-2), 124.60 (C-3

Please cite this article in press as: V. Desplat, et al., Synthesis and evaluation of the cytotoxic activity of novel ethyl 4-[4-(4-substitutedpiperidin-1-yl)]benzyl-phenylpyrrolo[1,2-*a*]quinoxaline-carboxylate derivatives in myeloid and lymphoid leukemia cell lines, European Journal of Medicinal Chemistry (2016), http://dx.doi.org/10.1016/j.ejmech.2016.02.047

pyr.), 123.69 (C-3), 121.79 (C-5 pyr.), 113.78 (C-9), 113.32 (C-1), 63.02 (CH_2N), 61.19 (CH_2), 53.39 (CH pip. and NCH_2), 30.74 (CH_2 pip.), 13.68 (CH_3). HRMS-ESI *m/z* $[M+Na]^+$ Calcd for $C_{39}H_{35}N_7O_2Na$: 656.2750, Found: 656.2768.

*5.1.10.4. Ethyl 4-{4-[(4-(2-oxo-2,3-dihydro-1H-benzimidazol-1-yl) piperidin-1-yl)benzyl]}-3-phenylpyrrolo[1,2-a]quinoxaline-2-carboxylate (**1d**).* White crystals (60%), mp 259–262 °C. ^{1}H NMR ($CDCl_3$) δ: 9.95 (s, 1H, NH), 8.67 (s, 1H, H-1), 8.07 (d, J = 8.0 Hz, 1H, H-6), 8.03 (d, J = 8.0 Hz, 1H, H-9), 7.61 (t, J = 8.0 Hz, 1H, H-7), 7.55 (t, J = 8.0 Hz, 1H, H-8), 7.34 (d, J = 7.2 Hz, 1H, H-4 benzimid.), 7.23 (d, J = 7.8 Hz, 2H, H-2′ and H-6′), 7.15–7.07 (m, 3H, H-5 benzimid., H-6 benzimid., H-7 benzimid.), 7.06–6.96 (m, 7H, H-3′, C-5′, H-2 phenyl, H-3 phenyl, H-4 phenyl, H-5 phenyl and H-6 phenyl), 4.45 (t, 1H, CH pip.), 4.23 (q, 2H, CH_2), 3.46 (s, 2H, CH_2N), 3.00 (d, J = 10.2 Hz, 2H, NCH_2 pip.), 2.52 (q, J = 7.2 Hz, 2H, CH_2 pip.), 2.15 (t, J = 11.4 Hz, 2H, NCH_2 pip.), 1.88 (d, J = 10.8 Hz, 2H, CH_2 pip.), 1.19 (t, J = 7.2 Hz, 3H, CH_3). ^{13}C NMR ($CDCl_3$) δ: 165.54 (C=O), 158.05 (C-4), 156.58 (C=O benzimid.), 140.33 (C-4′), 137.93 (C-1′), 137.56 (C-3a and C-5a), 135.22 (C-1 phenyl), 132.62 (C-3a benzimid., C-3 phenyl, C-4 phenyl and C-5 phenyl), 131.80 (C-6), 130.70 (C-7a benzimid.), 130.06 (2′ and C-6′), 129.54 (C-7, C-3′ and C-5′), 128.20 (C-2 phenyl and C-6 phenyl), 127.80 (C-8), 126.81 (C-9a), 124.41 (C-2), 122.61 (C-5 benzimid.), 122.52 (C-6 benzimid.), 120.66 (C-3), 119.65 (C-1), 115.21 (C-9) 111.25 (C-7 benzimid.), 111.17 (C-4 benzimid.), 64.03 (CH_2N), 61.79 (CH_2), 54.50 (NCH_2 pip.), 52.37 (CH pip.), 30.81 (CH_2 pip.), 15.53 (CH_3). HRMS-ESI *m/z* $[M+H]^+$ Calcd for $C_{39}H_{36}N_5O_3$: 622.2818, Found: 622.2795.

*5.1.10.5. Ethyl 4-{4-[(4-(5-fluoro-1H-benzimidazol-2-yl)piperidin-1-yl)benzyl]}-3-phenylpyrrolo[1,2-a]quinoxaline-2-carboxylate (**1e**).* White crystals (78%), mp 244–247 °C. ^{1}H NMR ($CDCl_3$) δ: 9.87 (s, 1H, NH), 8.67 (s, 1H, H-1), 8.04 (d, J = 7.0 Hz, 1H, H-9), 8.03 (t, J = 7.0 Hz, 1H, H-6), 7.61 (t, J = 7.0 Hz, 1H, H-7), 7.54 (t, J = 7.0 Hz, 1H, H-8), 7.18 (d, J = 8.0 Hz, 2H, H-2′ and H-6′), 7.06–6.96 (m, 7H, H-3′, H-5′, H-2 phenyl, H-3 phenyl, H-4 phenyl, H-5 phenyl, H-6 phenyl, H-4 benzimid., H-6 benzimid. and H-7 benzimid.), 4.23 (q, J = 7.2 Hz, 2H, CH_2), 3.41 (s, 2H, CH_2N), 2.97–2.86 (m, 3H, NCH_2 pip. and CH pip.), 2.11 (d, J = 11.6 Hz, 2H, CH_2 pip.), 2.08 (t, J = 11.6 Hz, 2H, NCH_2 pip.), 1.93 (m, 2H, CH_2 pip.), 1.20 (t, J = 7.2 Hz, 3H, CH_3). ^{13}C NMR ($CDCl_3$) δ: 165.56 (C=O), 161.10 (d, J = 236 Hz, C-5 benzimid.), 158.21 (C-4), 140.16 (C-4′), 138.00 (C-1 phenyl, C-1′, C-2 benzimid. and C-3a benzimid.), 137.57 (C-3a and C-5a), 135.19 (C-7a benzimid.), 131.62 (C-6), 130.11 (C-2′ and C-6′), 129.64 (3′ and C-5′), 129.54 (C-7), 128.19 (C-8), 128.04 (C-2 phenyl, C-3 phenyl, C-4 phenyl, C-5 phenyl and C-6 phenyl), 127.77 (9a), 124.44 (C-2), 120.69 (C-3), 119.69 (C-1), 115.31 (C-9 and C-7 benzimid.), 111.90 (C-4 benzimid. and C-6 benzimid.), 64.24 (CH_2N), 61.92 (CH_2), 54.66 (NCH_2 pip.), 38.34 (CH pip.), 32.34 (CH_2 pip.), 15.60 (CH_3). HRMS-ESI *m/z* $[M+H]^+$ Calcd for $C_{39}H_{35}N_5O_2F$: 624.2775, Found: 624.2783.

*5.1.10.6. Ethyl 4-{4-[4-(3-(pyridin-2-yl)-1,2,4-triazol-5-yl)piperidin-1-yl)benzyl]}-3-phenylpyrrolo[1,2-a]quinoxaline-2-carboxylate (**1f**).* Pale yellow crystals (53%), mp 138–140 °C. ^{1}H NMR ($CDCl_3$) δ: 8.73 (d, J = 3.0 Hz, 1H, H-6 pyr.), 8.64 (s, 1H, H-1), 8.23 (d, J = 7.8 Hz, 1H, H-5 pyr.), 8.02 (d, J = 8.4 Hz, 1H, H-9), 7.99 (d, J = 8.4 Hz, 1H, H-6), 7.84 (t, J = 7.8 Hz, 1H, H-4 pyr.), 7.56 (t, J = 7.8 Hz, 1H, H-7), 7.50 (t, J = 7.2 Hz, 1H, H-8), 7.36 (t, J = 6.0 Hz, 1H, H-3 pyr.), 7.17 (d, J = 7.2 Hz, 2H, H-2′ and H-6′), 7.06–7.02 (m, 1H, H-4 phenyl),7.01–6.94 (m, 6H, H-3′, H-5′, H-2 phenyl, H-3 phenyl, H-5 phenyl, and H-6 phenyl), 4.20 (q, J = 7.2 Hz, 2H, CH_2), 3.49 (s, 2H, CH_2N), 3.01–2.93 (m, 3H, NCH_2 pip. and CH pip.), 2.23–2.13 (m, 4H, NCH_2 pip. and CH_2 pip.), 2.12–2.03 (m, 2H, CH_2 pip.), 1.17 (t, J = 7.2 Hz, 3H, CH_3). ^{13}C NMR ($CDCl_3$) δ: 164.07 (C=O), 156.45 (C-4, C-3 Triazole and C-2 pyr.), 149.47 (C-6 pyr.), 147.36 (C-5 Triazole), 138.47 (C-4′ and C-4 pyr.), 136.63 (C-1 phenyl and C-1′), 135.97 (C-5a), 133.63 (C-3a), 130.21 (C-6), 128.64 (C-2′ and C-6′), 128.48 (C-3′ and C-5′), 128.09 (C-7), 126.77 (C-8), 126.49 (C-3 phenyl, C-4 phenyl and C-5 phenyl), 126.42 (C-2 phenyl and C-6 phenyl), 126.31 (9a), 125.41 (C-3 pyr.), 122.87 (C-2), 121.87 (C-5 pyr.), 119.12 (C-3), 118.26 (C-9), 113.76 (C-1), 62.64 (CH_2N), 60.34 (CH_2), 53.03 (NCH_2 pip.), 35.14 (CH pip.), 30.41 (CH_2 pip.), 14.07 (CH_3). HRMS-ESI *m/z* $[M+H]^+$ Calcd for $C_{39}H_{36}N_7O_2$: 634.2930, Found: 634.2952.

*5.1.10.7. Ethyl 4-{4-[(4-(2-oxo-2,3-dihydro-1H-benzimidazol-1-yl) piperidin-1-yl)benzyl]}-1-phenylpyrrolo[1,2-a]quinoxaline-3-carboxylate (**1g**).* Pale yellow crystals (66%), mp 167–169 °C. ^{1}H NMR ($CDCl_3$) δ: 10.52 (s, 1H, NH), 8.10 (d, J = 8.0 Hz, 1H, H-6), 7.82 (d, J = 7.6 Hz, 2H, H-2′ and H-6′), 7.62–7.53 (m, 8H, H-4 benzimid., H-2 phenyl, H-3 phenyl, H-4 phenyl, H-5 phenyl, H-6 phenyl, H-3′ and H-5′), 7.44 (d, J = 8.0 Hz, 1H, H-9), 7.43 (t, J = 8.0 Hz, 1H, H-7), 7.33 (d, J = 8.0 Hz, 1H, H-5 benzimid.), 7.21–7.02 (m, 4H, H-2, H-8, H-6 benzimid. and C-7 benzimid.), 4.52–4.36 (m, 1H, CH pip.), 3.79 (q, J = 7.2 Hz, 2H, CH_2), 3.69 (s, 2H, CH_2N), 3.13 (d, J = 10.8 Hz, 2H, NCH_2 pip.), 2.62–2.48 (m, 2H, CH_2 pip.), 2.27 (t, J = 10.8 Hz, 2H, NCH_2 pip.), 1.86 (d, J = 14.8 Hz, 2H, CH_2 pip.), 0.98 (t, J = 7.2 Hz, 3H, CH_3). ^{13}C NMR ($CDCl_3$) δ: 166.54 (C=O), 156.74 (C=O benzimid.), 155.97 (C-4), 141.21 (C-3a and C-4′), 140.77 (C-5a), 134.78 (C-1′), 133.44 (C-1 phenyl and C-3a benzimid.), 131.68 (C-6), 131.12 (C-3′ and C-5′), 130.56 (C-7a benzimid.), 130.37 (C-2 phenyl, C-3 phenyl, C-4 phenyl, C-5 phenyl and C-6 phenyl), 129.64 (C-2′ and C-6′), 129.46 (C-7), 128.66 (C-8), 127.22 (C-9a), 122.56 (C-1, C-3 and C-5 benzimid.), 120.22 (C-2 and C-6 benzimid.), 118.10 (C-9), 111.22 (C-7 benzimid.), 111.19 (C-4 benzimid.), 64.09 (CH_2N), 62.27 (CH_2), 54.60 (NCH_2 pip.), 52.18 (CH pip.), 30.69 (CH_2 pip.), 15.27 (CH_3). HRMS-ESI *m/z* $[M+H]^+$ Calcd for $C_{39}H_{36}N_5O_3$: 622.2818, Found: 622.2802.

*5.1.10.8. Ethyl 4-{4-[(4-(5-fluoro-1H-benzimidazol-2-yl)piperidin-1-yl)benzyl]}-1-phenylpyrrolo[1,2-a]quinoxaline-3-carboxylate (**1h**).* Pale yellow crystals (60%), mp 156–158 °C. ^{1}H NMR ($CDCl_3$) δ: 11.40 (s, 1H, NH), 8.02 (d, J = 7.6 Hz, 1H, H-6), 7.74 (d, J = 8.0 Hz, 2H, H-2′ and H-6′), 7.61–7.52 (m, 5H, H-2 phenyl, H-3 phenyl, H-4 phenyl, H-5 phenyl and H-6 phenyl), 7.46 (d, J = 8.0 Hz, 1H, H-9), 7.44–7.37 (m, 3H, H-7, H-3′ and H-5′), 7.21–7.15 (m, 4H, H-2, H-8, H-4 benzimid. and H-7 benzimid.), 6.96 (td, J = 9.2 Hz and 1.6 Hz, 1H, H-6 benzimid.), 3.72 (q, J = 7.2, 2H, CH_2), 3.55 (s, 2H, CH_2N), 2.98–2.84 (m, 3H, NCH_2 pip. and CH pip.), 2.13–1.98 (m, 4H, NCH_2 pip. and CH_2 pip.), 1.96–1.81 (m, 2H, CH_2 pip.), 0.92 (t, J = 7.2 Hz, 3H, CH_3). ^{13}C NMR ($CDCl_3$) δ: 166.61 (C=O), 160.7 (d, J = 235 Hz, C-5 benzimid.), 156.16 (C-4), 141.49 (C-4′), 140.40 (C-3a and C-5a), 138.66 (C-2 benzimid. and C-3a benzimid.), 134.65 (C-1 phenyl), 133.67 (C-7a benzimid.), 131.22 (C-3′ and C-5′), 131.11 (C-3 phenyl, C-4 phenyl and C-5 phenyl), 130.63 (C-6, C-2 phenyl and C-6 phenyl), 130.39 (C-7), 129.35 (C-2′ and C-6′), 128.72 (C-8), 128.54 (C-9a), 127.38 (C-1, C-3 and C-7 benzimid.), 120.33 (C-6 benzimid.), 118.21 (C-9), 111.50 (C-4 benzimid.), 64.19 (CH_2N), 62.32 (CH_2), 54.70 (NCH_2 pip.), 38.34 (CH pip.), 32.34 (CH_2 pip.), 15.21 (CH_3). HRMS-ESI *m/z* $[M+H]^+$ Calcd for $C_{39}H_{35}N_5O_2F$: 624.2775, Found: 624.2800.

*5.1.10.9. Ethyl 4-{4-[4-(3-(pyridin-2-yl)-1,2,4-triazol-5-yl)piperidin-1-yl)benzyl]}-1-phenylpyrrolo[1,2-a]quinoxaline-3-carboxylate (**1i**).* Pale yellow crystals (69%), mp 145–147 °C. ^{1}H NMR ($CDCl_3$) δ: 9.12 (s, 1H, NH), 8.68 (d, J = 4.0 Hz, 1H, H-6 pyr.), 8.20 (d, J = 8.0 Hz, 1H, H-5 pyr.), 8.06 (dd, J = 8.2 and 1.6 Hz, 1H, H-6), 7.86 (td, J = 7.6 and 1.6 Hz, 1H, H-4 pyr.), 7.78 (d, J = 7.4 Hz, 2H, H-2′ and H-6′), 7.57–7.54 (m, 5H, H-2 phenyl, H-3 phenyl, H-4 phenyl, H-5 phenyl and H-6 phenyl), 7.52 (d, J = 7.4 Hz, 2H, H-3′ and H-5′), 7.43 (d, J = 8.2 Hz, 1H, H-9), 7.41 (t, J = 7.2 Hz, 1H, H-7), 7.38 (dd, J = 6.4 and 1.2 Hz, 1H, H-9), 7.37 (d, J = 7.2 Hz, 1H, H-3 pyr.), 7.18 (td, J = 7.2 Hz and 1.6 Hz, 1H, H-8), 7.17 (s, 1H, H-2), 3.78 (q, J = 7.2, 2H, CH_2), 3.73

Please cite this article in press as: V. Desplat, et al., Synthesis and evaluation of the cytotoxic activity of novel ethyl 4-[4-(4-substitutedpiperidin-1-yl)]benzyl-phenylpyrrolo[1,2-*a*]quinoxaline-carboxylate derivatives in myeloid and lymphoid leukemia cell lines, European Journal of Medicinal Chemistry (2016), http://dx.doi.org/10.1016/j.ejmech.2016.02.047

(s, 2H, CH_2N), 3.14–3.12 (m, 2H, NCH_2 pip.), 3.00–2.88 (m, 1H, CH pip.), 2.37–2.30 (m, 2H, CH_2 pip.), 2.19–2.05 (m, 4H, 2 CH_2 pip.), 0.97 (t, J = 7.2 Hz, 3H, CH_3). ^{13}C NMR ($CDCl_3$) δ: 166.40 (C=O), 157.58 (C-2 pyr.), 155.85 (C-4), 148.75 (C-3 triazole), 150.77 (C-6 pyr.), 140.93 (C-5 triazole), 139.77 (C-3a), 138.84 (C-4 pyr.), 138.70 (C-5a), 134.70 (C-4′), 133.41 (C-1′ and C-1 phenyl), 131.59 (C-7), 131.06 (C-3′ and C-5′), 130.91 (C-3 phenyl and C-5 phenyl), 130.49 (C-8), 130.32 (C-2′ and C-6′), 129.46 (C-2 phenyl and C-6 phenyl), 128.61 (C-9a), 128.37 (C-6), 127.16 (C-9), 125.92 (C-4 phenyl), 125.77 (C-3), 123.23 (C-3 pyr.), 120.22 (C-5 pyr.), 118.04 (C-2), 115.92 (C-1), 64.02 (NCH_2), 62.24 (CH_2), 54.44 (NCH_2 pip.), 36.29 (CH pip.), 31.67 (CH_2 pip.), 15.22 (CH_3). HRMS-ESI *m/z* $[M + H]^+$ Calcd for $C_{39}H_{36}N_7O_2$: 634.2930, Found: 634.2959.

*5.1.10.10. Ethyl 4-{4-[(4-(2-oxo-2,3-dihydro-1H-benzimidazol-1-yl) piperidin-1-yl)benzyl]}-1-phenylpyrrolo[1,2-a]quinoxaline-2-carboxylate (**1j**).* White crystals (70%), mp 165–167 °C. ^{1}H NMR ($CDCl_3$) δ: 10.19 (s, 1H, NH), 8.03 (d, J = 7.8 Hz, 1H, H-6), 8.02 (d, J = 7.8 Hz, 2H, H-2′ and H-6′), 7.67–7.57 (m, 6H, H-7, H-3 phenyl, H-4 phenyl, H-5 phenyl, H-3′ and H-5′), 7.55 (s, 1H, H-3), 7.54–7.51 (m, 2H, H-2 phenyl and H-6 phenyl), 7.41–7.33 (m, 2H, H-9 and H-4 benzimid.), 7.14–7.05 (m, 4H, H-8, H-5 benzimid., H-6 benzimid., H-7 benzimid.), 4.47 (t, J = 12.0 Hz, 1H, CH pip.), 4.17 (q, J = 7.2 Hz, 2H, CH_2), 3.73 (s, 2H, CH_2N), 3.23–3.08 (m, 2H, NCH_2 pip.), 2.65–2.50 (m, 2H, CH_2 pip.), 2.37–2.23 (m, 2H, NCH_2 pip.), 1.89 (d, J = 10.2 Hz, 2H, CH_2 pip.), 1.13 (t, J = 7.2 Hz, 3H, CH_3). ^{13}C NMR ($CDCl_3$) δ: 165.59 (C=O), 156.75 (C-4), 156.60 (C=O benzimid.), 139.29 (C-5a and C-3a benzimid.), 138.22 (C-1′), 136.20 (C-3a and C-1 phenyl), 131.88 (C-6), 131.85 (C-2 phenyl and C-6 phenyl), 130.85 (C-7a benzimid.), 130.80 (C-4 phenyl), 130.51 (C-3′ and C-5′), 130.24 (C-7), 130.12 (C-3 phenyl and C-5 phenyl), 129.67 (C-2′ and C-6′), 128.19 (C-8), 127.34 (C-9), 126.61 (C-2), 120.81 (C-1), 118.12 (C-5 benzimid. and C-6 benzimid.), 111.94 (C-3 and C-4 benzimid.), 111.20 (C-7 benzimid.), 64.01 (CH_2N), 61.61 (CH_2), 54.61 (NCH_2 pip.), 52.15 (CH pip.), 30.67 (CH_2 pip.), 15.42 (CH_3). HRMS-ESI *m/z* $[M+H]^+$ Calcd for $C_{39}H_{36}N_5O_3$: 622.2818, Found: 622.2846.

*5.1.10.11. Ethyl 4-{4-[(4-(5-fluoro-1H-benzimidazol-2-yl)piperidin-1-yl)benzyl]}-1-phenylpyrrolo[1,2-a]quinoxaline-2-carboxylate (**1k**).* White crystals (46%), mp 195–197 °C. ^{1}H NMR ($CDCl_3$) δ: 7.97 (d, J = 7.8 Hz, 1H, H-6), 7.93 (d, J = 7.2 Hz, 2H, H-2′ and H-6′), 7.64–7.56 (m, 3H, H-7, H-9 and H-4 phenyl), 7.55–7.47 (m, 6H, H-3, H-3 phenyl, H-5 phenyl, H-3′, H-5′ and H-4 benzimid.), 7.36 (t, J = 6.6 Hz, 2H, H-2 phenyl and H-6 phenyl), 7.09 (d, J = 7.8 Hz, 1H, H-8), 7.08 (d, J = 7.8 Hz, 1H, H-7 benzimid.), 6.98 (t, J = 7.8 Hz, 1H, H-6 benzimid.), 4.16 (q, J = 7.2 Hz, 2H, CH_2), 3.62 (s, 2H, CH_2N), 2.99 (d, J = 11.4 Hz, 2H, NCH_2 pip.), 2.94 (t, J = 12 Hz, 1H, CH pip.), 2.15 (t, J = 11.4 Hz, 2H, NCH_2 pip.), 2.06 (d, J = 11.4 Hz, 2H, CH_2 pip.), 1.98–1.89 (m, 2H, CH_2 pip.), 1.11 (t, J = 7.2 Hz, 3H, CH_3). ^{13}C NMR ($CDCl_3$) δ: 165.75 (C=O), 160.57 (d, J = 235 Hz, C-5 benzimid.), 156.90 (C-4), 142.10 (C-4′), 139.21 (C-5a), 137.97 (C-1′), 136.42 (C-2 benzimid. and C-3a benzimid.), 134.60 (C-1 phenyl), 132.51(C-9a), 131.90 (C-3′ and C-5′), 131.52 (C-3a), 130.91 (C-6), 130.84 (C-3 phenyl and C-5 phenyl), 130.32 (C-7a benzimid.), 130.05 (C-7, C-9 and C-4 phenyl.), 129.58 (C-2′ and C-6′), 128.37 (C-8), 127.52 (C-2 phenyl and C-6 phenyl), 120.93 (C-1 and C-2), 118.28 (C-7 benzimid.), 111.82 (C-3 and C-6 benzimid.), 111.64 (C-4 benzimid.), 64.24 (CH_2N), 61.79 (CH_2), 54.74 (NCH_2 pip.), 38.31 (CH pip.), 32.35 (CH_2 pip.), 15.41 (CH_3). HRMS-ESI *m/z* $[M+H]^+$ Calcd for $C_{39}H_{35}N_5O_2F$: 624.2775, Found: 624.2747.

*5.1.10.12. Ethyl 4-{4-[4-(3-(pyridin-2-yl)-1,2,4-triazol-5-yl)piperidin-1-yl)benzyl]}-1-phenylpyrrolo[1,2-a]quinoxaline-2-carboxylate (**1l**).* White crystals (42%), mp 144–146 °C. ^{1}H NMR ($CDCl_3$) δ: 8.69 (d, J = 4.6 Hz, 1H, H-6 pyr.), 8.19 (d, J = 7.8 Hz, 1H, H-5 pyr.), 7.99 (d, J = 7.0 Hz, 1H, H-6), 7.94 (d, J = 7.0 Hz, 2H, H-2′ and H-6′), 7.83 (td, J = 7.6 and 1.6 Hz, 1H, H-4 pyr.), 7.62–7.45 (m, 8H, H-3, H-7, H-9, H-3 phenyl, H-4 phenyl, H-5 phenyl, H-3′ and H-5′), 7.40–7.29 (m, 3H, H-2 phenyl, H-6 phenyl and H-3 pyr.), 7.04 (d, J = 7.0 Hz, 1H, H-8), 4.14 (q, J = 7.1 Hz, 2H, CH_2), 3.67 (s, 2H, CH_2N), 3.07 (d, J = 11.0 Hz, 2H, NCH_2 pip.), 2.99–2.79 (m, 1H, CH pip.), 2.34–1.90 (m, 6H, NCH_2 pip. and 2 × CH_2 pip.), 1.10 (t, J = 7.1 Hz, 3H, CH_3). ^{13}C NMR ($CDCl_3$) δ: 165.65 (C=O), 156.79 (C-4 and C-2 pyr.), 150.99 (C-3 Triazole and C-6 pyr.), 148.41 (C-5 Triazole), 139.39 (C-4′), 138.86 (C-5a and C-4 pyr.), 138.18 (C-1′), 134.72 (C-1 phenyl), 131.94 (C-3a, C-3′ and C-5′), 130.98 (C-6), 130.86 (C-3 phenyl and C-5 phenyl), 130.11 (C-7, C-9 and C-4 phenyl), 129.56 (C-2′ and C-6′), 128.22 (C-8), 127.37 (C-3 pyr., C-2 phenyl and C-6 phenyl), 123.11 (C-2, C-9a and C-5 pyr.), 120.84 (C-1), 112.01 (C-3), 64.51 (CH_2N), 61.68 (CH_2), 55.03 (NCH_2 pip.), 32.32 (CH_2 pip. and CH pip.), 15.49 (CH_3). HRMS-ESI *m/z* $[M + H]^+$ Calcd for $C_{39}H_{36}N_7O_2$: 634.2930, Found: 634.2962.

5.2. X-ray data

The structure of compounds **2b-d**, **1a**, **1d**, **1g**, **1j** and **1k** has been established by X-ray crystallography (Figs. 3 and 4). Colorless single crystal of **2b** was obtained by slow evaporation from chloroform: monoclinic, space group $P2_1/c$, a = 11.9390(11) Å, b = 13.8893(9) Å, c = 18.5001(19) Å, α = 90°, β = 100.959(7)°, γ = 90°, V = 3011.8(5) Å^3, Z = 8, δ(calcd) = 1.267 Mg m^{-3}, FW = 287.31 for $C_{16}H_{17}NO_4$, $F(000)$ = 1216. Colorless single crystal of **2c** was obtained by slow evaporation from methanol/dichloromethane (20/80) solution: monoclinic, space group $P2_1/n$, a = 14.653(3) Å, b = 5.1458(12) Å, c = 18.842(4) Å, α = 90°, β = 94.447(14)°, γ = 90°, V = 1416.4(5) Å^3, Z = 4, δ(calcd) = 1.347 Mg m^{-3}, FW = 287.31 for $C_{16}H_{17}NO_4$, $F(000)$ = 608. Colorless single crystal of **2d** was obtained by slow evaporation from methanol/dichloromethane (20/80) solution: triclinic, space group P-1, a = 8.628(9) Å, b = 9.798(2) Å, c = 10.095(3) Å, α = 73.02(2)°, β = 71.00(3)°, γ = 73.95(5)°, V = 756.1(8) Å^3, Z = 1, δ(calcd) = 1.262 Mg m^{-3}, FW = 574.61 for $C_{32}H_{34}N_2O_8$, $F(000)$ = 304. Colorless single crystal of **1a** was obtained by slow evaporation from methanol/dichloromethane (30/70) solution: triclinic, space group P-1, a = 10.0140(10) Å, b = 10.877(2) Å, c = 17.580(2) Å, α = 79.642(9)°, β = 86.041(8)°, γ = 65.016(8)°, V = 1707.3(4) Å^3, Z = 2, δ(calcd) = 1.311 Mg m^{-3}, FW = 673.96 for $C_{39}H_{35}N_5O_3$,0.48($CH_2Cl_2O_2$),0.19(O), $F(000)$ = 710. Colorless single crystal of **1d** was obtained by slow evaporation from methanol/dichloromethane (30/70) solution: triclinic, space group P-1, a = 8.3759(10) Å, b = 12.6418(13) Å, c = 16.6718(18) Å, α = 109.502(7)°, β = 96.482(8)°, γ = 103.928(7)°, V = 1578.7(3) Å^3, Z = 2, δ(calcd) = 1.308 Mg m^{-3}, FW = 621.72 for $C_{39}H_{35}N_5O_3$, $F(000)$ = 656. Colorless single crystal of **1g** was obtained by slow evaporation from methanol/dichloromethane (20/80) solution: triclinic, space group P-1, a = 9.6882(14) Å, b = 11.2689(13) Å, c = 19.639(2) Å, α = 77.504(8)°, β = 83.067(11)°, γ = 72.264(11)°, V = 1990.4(4) Å^3, Z = 2, δ(calcd) = 1.347 Mg m^{-3}, FW = 807.16 for $C_{39}H_{35}N_5O_3$,1.56($CHCl_3$), $F(000)$ = 836. Colorless single crystal of **1j** was obtained by slow evaporation from methanol/dichloromethane (30/70) solution: triclinic, space group P-1, a = 9.7518(8) Å, b = 11.1004(8) Å, c = 17.9247(12) Å, α = 86.622(5)°, β = 82.713(5)°, γ = 66.141(5)°, V = 1760.1(2) Å^3, Z = 2, δ(calcd) = 1.398 Mg m^{-3}, FW = 741.09 for $C_{39}H_{35}N_5O_3$,$CHCl_3$, $F(000)$ = 772. Pale-yellow single crystal of **1k** was obtained by slow evaporation from methanol/dichloromethane (30/70) solution: monoclinic, space group C2/c, a = 18.0617(14) Å, b = 9.4210(8) Å, c = 43.021(3) Å, α = 90°, β = 93.345(5)°, γ = 90°, V = 7307.9(10) Å^3, Z = 8, δ(calcd) = 1.296 Mg m^{-3}, FW = 712.78 for $C_{39}H_{34}FN_5O_2$,5(H_2O) $F(000)$ = 3016. Full crystallographic results have been deposited at the Cambridge Crystallographic Data Centre (CCDC-1014944, CCDC-891817, CCDC-891816, CCDC-891812, CCDC-

Please cite this article in press as: V. Desplat, et al., Synthesis and evaluation of the cytotoxic activity of novel ethyl 4-[4-(4-substitutedpiperidin-1-yl)]benzyl-phenylpyrrolo[1,2-*a*]quinoxaline-carboxylate derivatives in myeloid and lymphoid leukemia cell lines, European Journal of Medicinal Chemistry (2016), http://dx.doi.org/10.1016/j.ejmech.2016.02.047

891811, CCDC-891813, CCDC-891814, CCDC-891815, respectively), UK, as Supplementary Material [38]. The data were corrected for Lorentz and polarization effects and for empirical absorption correction [39]. The structure was solved by direct methods Shelx 2013 [40] and refined using Shelx 2013 [40] suite of programs.

5.3. Biology

5.3.1. Cell culture

The human leukemic cell lines U937, K562, HL60, U266 and Jurkat were grown in RPMI 1640 medium (Life Technology, France) supplemented with 10% fetal calf serum (FCS), antibiotics (100 U/mL penicillin, 100 μg/mL streptomycin) and L-glutamin, (Eurobio, France) at 37 °C, 5% CO_2 in air. The toxicity of various molecules was also evaluated on non-activated, freshly isolated normal human peripheral blood mononuclear cells (PBMNC), as well as phytohemagglutinin (T lymphoproliferative agent) (PHA)-induced cells. PBMNC from blood of healthy volunteers were obtained following centrifugation on Ficoll gradient. Cells were then incubated in medium alone or induced to enter cell cycle by the addition of PHA (5 μg/mL, Murex Biotech Limited, Dartford, UK).

5.3.2. Cytotoxicity test

The MTS cell proliferation assay (Promega, France) is a colorimetric assay system, which measures the reduction of a tetrazolium component (MTS) into formazan produced by the mitochondria of viable cells. Cells were washed twice in PBS (Phosphate Buffer Saline) and plated in quadruplicate into microtiter-plate wells in 100 μL culture media with or without our various compounds at increasing concentrations (0, 1, 5, 10, 20 and 50 μM) during 1, 2 and 3 days. After 3 h of incubation at 37 °C with 20 μL MTS/well, the plates were read by using an ELISA microplate reader (Thermo, Electrocorporation) at 490 nm wavelength. The amount of colour produced was directly proportional to the number of viable cells. The results are expressed as the concentrations inhibiting cell growth by 50% after a 3 days incubation period. The 50% cytotoxic concentrations (CC_{50}) were determined by linear regression analysis, expressed in μM ± SD (Microsoft Excel).

5.3.3. Annexin V staining by flow cytometry

Cells (2×10^5) were incubated for 3 days with increasing doses of **1e** (0, 1, 5, and 10). Experiments were performed with APC-Annexin V (Biolegend, CA) according to the manufacturer's instructions. Briefly, cells (2×10^4) were incubated with 5 mL of APC-Annexin V resuspended in 295 mL of 1X binding buffer for 10 min at room temperature in the dark. Then, cells were analyzed by flow cytometry. The APC-Annexin V-positive cells were considered as apoptotic. Flow cytometry analysis was performed with a BD FACS Canto II flow cytometer (Becton-Dickinson, France) and experiments were analyzed using the Diva software.

5.3.4. Computational prediction of toxicity and drug relevant properties

Calculations of Clog *P* and calculations of Topological Polar Surface Area (TPSA), number of hydrogen bond acceptor (nON) and donor (nNH/OH) atoms, and any violations to the Lipinski's "rule of five" (logP ≤ 5, molecular weight ≤500, number of hydrogen bond acceptors ≤10, and number of hydrogen bond donors ≤5) [34]; were performed using the MIPC server at http://www.molinspiration.com/cgi-bin/properties [33].

Acknowledgments

This publication was supported by a grant from Ligue Nationale contre le Cancer (Comité Aquitaine-Charentes, Bordeaux, France). Marian Vincenzi thanks the Università Italo Francese (UIF) for financial support (project C2-1, 2014/38723, cap. 6.01.1810 UIF).

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.ejmech.2016.02.047.

References

[1] M.A. Lichtman, Battling the hematological malignancies: the 200 years' war, The Oncologist 13 (2008) 126–138.
[2] G. Campiani, S. Butini, C. Fattorusso, F. Trotta, S. Franceschina, M. De Angelis, K.S. Nielsen, Novel Aryl Piperazine Derivatives with Medical Utility, 2006. WO2006072608.
[3] G. Campiani, F. Aiello, M. Fabbrini, E. Morelli, A. Ramunno, S. Armaroli, V. Nacci, A. Garofalo, G. Greco, E. Novellino, G. Maga, S. Spadari, A. Bergamini, L. Ventura, B. Bongiovanni, M. Capozzi, F. Bolacchi, S. Marini, M. Coletta, G. Guiso, S. Caccia, Quinoxalinylethylpyridylthioureas (QXPTs) as potent non-nucleoside HIV-1 reverse transcriptase (RT) inhibitors. Further SAR studies and identification of a novel orally bioavailable hydrazine-based antiviral agent, J. Med. Chem. 44 (2001) 305–315.
[4] S. Schann, S. Mayer, S. Gardan, Pyrrolo[1,2-*a*]quinoxaline Derivatives as Adenosine A3 Receptor Modulators and Uses Thereof, 2007. EP1798233.
[5] J. Guillon, P. Grellier, M. Labaied, P. Sonnet, J.-M. Léger, R. Déprez-Poulain, I. Forfar-Bares, P. Dallemagne, N. Lemaître, F. Péhourcq, J. Rochette, C. Sergheraert, C. Jarry, Synthesis, antimalarial activity and molecular modeling of new pyrrolo[1,2-*a*]quinoxalines, bispyrrolo[1,2-*a*]quinoxalines, bispyrido[3,2-*e*]pyrrolo[1,2-*a*]pyrazines and bispyrrolo[1,2-*a*]thieno[3,2-*e*]pyrazines, J. Med. Chem. 47 (2004) 1997–2009.
[6] J. Guillon, I. Forfar, M. Mamani-Matsuda, V. Desplat, M. Saliège, D. Thiolat, S. Massip, A. Tabourier, J.-M. Léger, B. Dufaure, G. Haumont, C. Jarry, D. Mossalayi, Synthesis, analytical behaviour and biological evaluation of new 4-substituted pyrrolo[1,2-*a*]quinoxalines as antileishmanial agents, Bioorg. Med. Chem. 15 (2007) 194–210.
[7] J. Guillon, I. Forfar, V. Desplat, S. Belisle-Fabre, D. Thiolat, S. Massip, H. Carrie, D. Mossalayi, C. Jarry, Synthesis of new 4-(*E*)-alkenylpyrrolo[1,2-*a*]quinoxalines as antileishmanial agents by suzuki-miyaura cross-coupling reactions, J. Enzyme Inhib. Med. Chem. 22 (2007) 541–549.
[8] J. Guillon, S. Moreau, E. Mouray, V. Sinou, I. Forfar, S. Belisle-Fabre, V. Desplat, P. Millet, D. Parzy, C. Jarry, P. Grellier, New ferrocenic pyrrolo[1,2-*a*]quinoxaline derivatives : synthesis, and in vitro antimalarial activity, Bioorg. Med. Chem. 16 (2008) 9133–9144.
[9] J. Guillon, E. Mouray, S. Moreau, C. Mullié, I. Forfar, V. Desplat, S. Belisle-Fabre, F. Ravanello, A. Le-Naour, N. Pinaud, J.-M. Léger, G. Gosmann, C. Jarry, G. Déléris, P. Sonnet, P. Grellier, New ferrocenic pyrrolo[1,2-*a*]quinoxaline derivatives: Synthesis, and *in vitro* antimalarial activity – part II, Eur. J. Med. Chem. 46 (2011) 2310–2326.
[10] L. Ronga, M. Del Favero, A. Cohen, C. Soum, P. Le Pape, S. Savrimoutou, N. Pinaud, C. Mullié, S. Daulouede, P. Vincendeau, N. Farvacques, P. Agnamey, F. Pagniez, S. Hutter, N. Azas, P. Sonnet, J. Guillon, Design, synthesis and biological evaluation of novel 4-Alkapolyenylpyrrolo[1,2-a]quinoxalines as antileishmanial agents – part III, Eur. J. Med. Chem. 81 (2014) 378–393.
[11] J. Milne, K.D. Normington, M. Milburn, Tetrahydroquinoxalinone Sirtuin Modulators, 2006. WO2006094210.
[12] F. Grande, F. Aiello, O. De Grazia, A. Brizzi, A. Garofalo, N. Meamati, Synthesis and antitumor activities of a series of novel quinoxalinhydrazides, Bioorg. Med. Chem. 15 (2007) 288–294.
[13] J. Guillon, M. Le Borgne, C. Rimbault, S. Moreau, S. Savrimoutou, N. Pinaud, S. Baratin, M. Marchivie, S. Roche, A. Bollacke, A. Pecci, L. Alvarez, V. Desplat, J. Jose, Synthesis and biological evaluation of novel substituted pyrrolo[1,2-*a*] quinoxaline derivatives as inhibitors of the human protein kinase CK2, Eur. J. Med. Chem. 65 (2013) 205–222.
[14] V. Desplat, A. Geneste, M.-A. Begorre, S. Belisle Fabre, S. Brajot, S. Massip, D. Thiolat, D. Mossalayi, C. Jarry, J. Guillon, Synthesis of new pyrrolo[1,2-*a*] quinoxaline derivatives as potential inhibitors of Akt kinase, J. Enzyme Inhib. Med. Chem. 23 (2008) 648–658.
[15] V. Desplat, S. Moreau, A. Gay, S. Belisle-Fabre, D. Thiolat, S. Massip, D. Mossalayi, C. Jarry, J. Guillon, Synthesis and evaluation of the anti-proliferative activity of novel pyrrolo[1,2-*a*]quinoxaline derivatives, potential inhibitors of akt kinase. Part II, J. Enzyme Inhib. Med. Chem. 25 (2010) 204–215.
[16] V. Desplat, S. Moreau, S. Belisle-Fabre, D. Thiolat, J. Uranga, R. Lucas, S. Massip, L. de Moor, C. Jarry, D. Mossalayi, P. Sonnet, G. Déléris, J. Guillon, Synthesis and evaluation of the antiproliferative activity of novel isoindolo[2,1-*a*]quinoxaline and Indolo[1,2-*a*]quinoxaline derivatives, J. Enzyme Inhib. Med. Chem. 26 (2011) 657–667.
[17] H.A. Jensen, R.P. Bunaciu, J.D. Varner, A. Yen, GW5074 and PP2 kinase inhibitors implicate nontraditional c-Raf and Lyn function as drivers of retinoic acid-induced maturation, Cell. Signal 27 (2015) 1666–1675.
[18] S. Kamijo, C. Kanazawa, Y. Yamamoto, Copper- or phosphino-catalyzed reaction of alkynes with isocyanides. Regioselective synthesis of substituted

pyrroles controlled by the catalyst, J. Am. Chem. Soc. 127 (2005) 9260–9266.
[19] A.V. Gulevich, A.G. Zhdanko, R.V.A. Orru, V.G. Nenajdenko, Isocyanoacetate derivatives: synthesis, reactivity, and application, Chem. Rev. 110 (2010) 5235–5331.
[20] J. Liu, Z. Fang, Q. Zhang, Q. Liu, X. Bi, Silver-catalyzed isocyanide-alkyne cycloaddition: a general and practical method to oligosubstituted pyrroles, Angew. Chem. Int. Ed. 52 (2013) 6953–6957.
[21] M. Gao, C. He, H. Chen, R. Bai, B. Cheng, A. Lei, Synthesis of pyrroles by click reaction: silver-catalyzed cycloaddition of terminal alkynes with isocyanides, Angew. Chem. Int. Ed. 52 (2013) 6958–6961.
[22] M. Suzuki, M. Miyoshi, K. Matsumoto, Synthesis of amino acids and related compounds. 7. A convenient synthesis of 3-substituted pyrrole-2,4-dicarboxylic acid esters, J. Org. Chem. 39 (1974) 1980.
[23] J. Guillon, H. Dumoulin, P. Dallemagne, R. Reynolds, S. Rault, Synthesis and antituberculosis activity of new phenylpyrrolo[1,2-*a*]quinoxalinylpyrrole carboxylic acid derivatives, Pharm. Pharmacol. Commun. 4 (1998) 33–38.
[24] S. Ngwerume, J.E. Camp, Synthesis of highly substituted pyrroles via nucleophilic catalysis, J. Org. Chem. 75 (2010) 6271–6274.
[25] B.A. Trofimov, Preparation of pyrroles from ketoximes and acetylenes, Adv. Heterocycl. Chem. 51 (1999) 177–301.
[26] Y.-F. Wang, K.K. Toh, S. Chiba, K. Narasaka, Mn(III)-catalyzed synthesis of pyrroles from vinyl azides and 1,3-dicarbonyl compounds, Org. Lett. 10 (2008) 5019–5022.
[27] P. Liu, B. Xu, J. Li, H. Lu, LY294002 inhibits leukemia cell invasion and migration through early growth response gene 1 induction independent of phosphatidylinositol 3-kinase-Akt pathway, Biochem. Biophys. Res. Commun. 377 (2008) 187–190.
[28] S. Matsuda, A. Nakanishi, Y. Wada, Y. Kitagishi, Roles of PI3K/AKT/PTEN pathway as a target for pharmaceutical therapy, Open Med. Chem. J. 7 (2013) 23–29.
[29] H.G. Jørgensen, E.K. Allan, S.M. Graham, J.L. Godden, L. Richmond, M.A. Elliott, J.C. Mountford, C.J. Eaves, T.L. Holyoake, Lonafarnib reduces the resistance of primitive quiescent CML cells to imatinib mesylate *in vitro*, Leukemia 19 (2005) 1184–1191.
[30] S. Zhao, M. Konopleva, M. Cabreira-Hansen, Z. Xie, W. Hu, M. Milella, Z. Estrov, G.B. Mills, M. Andreeff, Inhibition of phosphatidylinositol 3-kinase dephosphorylates BAD and promotes apoptosis in myeloid leukemias, Leukemia 18 (2004) 267–275.
[31] C.W. Lindsley, The Akt/PKB family of protein kinases: a review of small molecule inhibitors and progress towards target validation: a 2009 update, Curr. Top. Med. Chem. 10 (2010) 458–477.
[32] R. Boulahjar, A. Ouach, M. Chiurato, S. Bourg, M. Ravache, R. Le Guével, S. Marionneau-Lambot, T. Oullier, O. Lozach, L. Meijer, C. Guguen-Guillouzo, S. Lazar, M. Akssira, Y. Troin, G. Guillaumet, S. Routier, Novel tetrahydropyrido [1,2-*a*]isoindolone derivatives (Valmerins): potent cyclin dependant kinase/glycogen synthase kinase 3 inhibitors with antiproliferative activities and antitumor effects in human tumor xenografts, J. Med. Chem. 55 (2012) 9589–9606.
[33] http://www.molinspiration.com/cgi-bin/properties. Accession date: July 20, 2015.
[34] D.E. Clark, S.D. Pickett, Computational methods for the prediction of drug likeness, Drug Discov. Today 5 (2000) 49–58.
[35] M.F.C. Cardoso, P.C. Rodrigues, M.E.I.M. Oliveira, I.L. Gama, I.M.C.B. da Silva, I.O. Santos, D.R. Rocha, R.T. Pinho, V.F. Ferreira, M.C.B.V. de Souza, F. de C. da Silva, F. Paes Silva Jr., Synthesis and evaluation of the cytotoxic activity of 1,2-furanonaphthoquinones tethered to 1,2,3-1H-triazoles in myeloid and lymphoid leukemia cell lines, Eur. J. Med. Chem. 84 (2014) 708–717.
[36] J.F. Bickley, T.L. Gilchrist, R. Mendonça, Hydroxylation of 1-azabicyclo[4.1.0] hept-3-enes formed by Diels-Alder reactions of benzyl 2*H*-azirine-3-carboxylate, Arkivoc vi (2002) 192–204.
[37] T.L. Gilchrist, R. Mendonça, Addition of pyrimidine and purine bases to benzyl 2*H*-azirine-3-carboxylate, Synlett 12 (2000) 1843–1845.
[38] Supplementary X-ray Crystallographic Data: Cambridge Crystallographic Data Centre, University Chemical Lab, Lensfield Road, Cambridge, CB2 1EW, UK; E-mail: deposit@chemcrys.cam.ac.uk.
[39] G.M. Sheldrick, SADABS, University of Göttingen, Germany, 1996.
[40] G.M. Sheldrick, Acta Crystallogr. Sect. A 64 (2008) 112–122.

Please cite this article in press as: V. Desplat, et al., Synthesis and evaluation of the cytotoxic activity of novel ethyl 4-[4-(4-substitutedpiperidin-1-yl)]benzyl-phenylpyrrolo[1,2-*a*]quinoxaline-carboxylate derivatives in myeloid and lymphoid leukemia cell lines, European Journal of Medicinal Chemistry (2016), http://dx.doi.org/10.1016/j.ejmech.2016.02.047

INTERACTIONS PROTÉINES-MOLÉCULES BIOTECHNOLOGIQUES ORIGINALES: UNE APPROCHE INTÉGRÉE DE RMN ET DE MODÉLISATION MOLÉCULAIRE

Résumé :

Il est connu que le processus de recherche et développement de nouveaux médicaments nécessite de nombreuses étapes. Ceux-ci comprennent, en plus de l'identification de la cible biologique, la synthèse, le criblage et la construction ultérieure des bibliothèques de molécules potentiellement dotésune activité pharmacologique. L'ensemble du processus est très exigeant en termes de coût (1-2 milliards d'euros) et de temps (12-13 ans) [1] En conséquence, les méthodes de la chimie numérique et la dynamique moléculaire ont le potentiel de contribuer au développement d'un médicament en termes de conception, rationalisation des relations structure-activité et donc de réduction des coûts. Dans ce contexte, la relation quantitative structure à activité (RSA), l'analyse de docking moléculaire et la dynamique moléculaire, peuvent contribuer efficacement à une rationalisation des résultats obtenus expérimentalement. L'objectif principal de cette approche multidisciplinaire vise à déterminer les détails conformationnels fondamentales pour l'interaction entre la protéine et son ligand, afin de avoir des informations utiles pour la conception de nouveaux médicaments potentiels. Les molécules identifiées par cette étude toujours besoin d'autres études expérimentales in vitro et in vivo pour avoir une compréhension approfondie du mécanisme d'action du médicament potentiel, et pour régler les problèmes qui influent fortement sur la pharmacocinétique du composé ou de ses métabolites (solubilité, la perméabilité des molécules et des interactions possibles avec d'autres biomolécules ou d'autres médicaments). Afin de définir le mécanisme d'action, l'analyse au niveau moléculaire des détails structuraux des interactions entre biomolécules permettre l'acquisition d'informations utiles pour leur compréhension et, par conséquent, pour la détermination des fonctionnalisations nécessaires pour augmenter l'activité du médicament potentiel. Dans ce contexte nous trouvons cette thèse qui concerne l'application d'une approche intégrée caractérisée par l'utilisation de différentes techniques spectroscopiques (fluorescence, dichroïsme circulaire (CD), résonance magnétique nucléaire (RMN) et diffusion dynamique de la lumière (DLS)) et de calculs (docking moléculaire, dynamique moléculaire) pour obtenir une meilleure compréhension des détails structurels en solution et des interactions intermoléculaires impliquées dans les mécanismes d'action de deux systèmes: Akt-ligand hétérocyclique (**Workpackage 1**) et la chimiokine, CXCR4 (**Workpackage 2**).

Workpackage 1

Le **Workpackage 1** de ma thèse concerne la conception et la synthèse de composés anti-prolifératifs nouveaux. Nous avons concentré sur les molécules capables de bloquer l'activité de l'Akt, dont la surexpression a été associée à une prolifération incontrôlée des cellules cancéreuses. Dans ce contexte, un travail récent met en évidence l'importance de l'anneau le pyrrolo[1,2-a]quinoxaline comme un échafaudage moléculaire et mettre en évidence le rôle vital des substitutions appropriées dans des

positions spécifiques pour moduler l'activité de la molécule entière comme agent antiprolifératif. De fait, la présence d'un cycle phényle et un groupe ester d'éthyle sur la partie pyrrole de cet échafaudage se sont révélés être importants pour l'activité de la molécule, comme le montrent des essais effectués sur différentes lignées de cellules tumorales humaines [5]. Ces études ont également montré l'influence significative des substituants benzylpiperidinylbenzimidazolone et fluorobenzimidazole sur l'activité anti-prolifération. Dans la littérature, il est rapporté que, d'un point de vue synthétique, la formation d'un noyau phényl-1H-pyrrol-diester constitue l'étape essentielle du processus entier. Pour cela, différents protocoles de synthèse déjà décrites dans la littérature pour des molécules similaires, ont été comparées, afin d'évaluer les effets des conditions différentes (température, solvant et catalyseur) sur les temps de réaction et le rendement. L'optimisation des conditions de synthèse a permis d'obtenir des phényl-1H-pyrrole-diesters, à partir de laquelle les échafauds pyrrolo [1,2-a] quinoxaline remplacé autrement (intermédiaire clé dans la synthèse des composés finaux **1a-l**) ont été synthétisés. Les structures sont résumées dans la **Figure 1**.

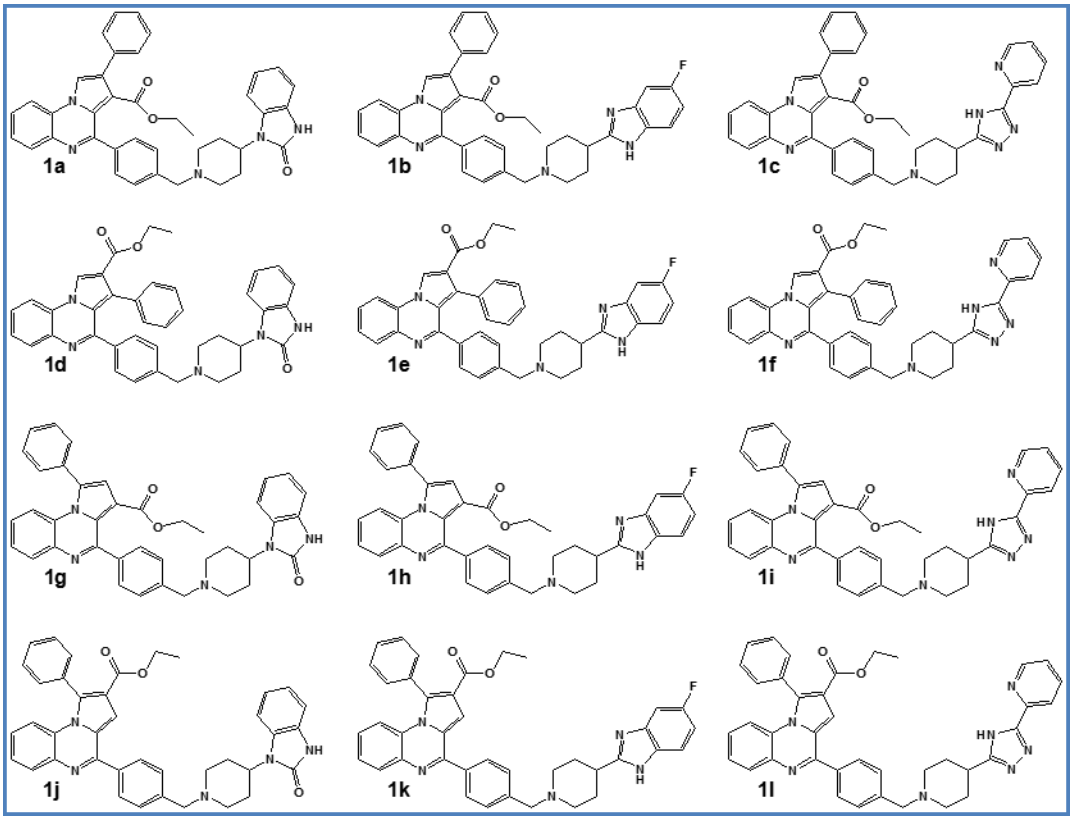


***Figure 1**. Structure chimique des composés **1a-l**.*

La structure des composés a été confirmée par la caractérisation IR et RMN. Tests d'activité anti-prolifération (MTS) ont également été menées sur différentes lignées de cellules cancéreuses humaines (U937, K562, Jurkat, U266 et HL60) afin de comparer les différents types de cellules tumorales dans lesquelles la protéine Akt, normalement impliqué dans la régulation du processus de prolifération, est surexprimé. Les molécules **1a**, **1e**, **1g** et **1h** ont montré une forte activité cytotoxique contre certaines lignées de cellules cancéreuses (IC_{50} dans la gamme 3-9 μM) et une faible toxicité contre les cellules hématopoïétiques normales (IC_{50}> 50 μM) (**Tableau 1**). Ces composés, montrant un indice de sélectivité intéressante, peuvent être des candidats appropriés pour des études pharmacologiques ultérieures. D'un point de vue général de la RAS, ces résultats biologiques préliminaires soulignent l'importance de la

présence d'un groupe benzylpiperidinylfluorobenzimidazole dans la position C-4 de l'échafaud pyrrole [1,2-a] quinoxaline, mais aussi la nécessité d'une fonctionnalisation sur l'anneau pyrrole. Les données biologiques semblent indiquer que le composé **1e** est le plus prometteur. Il serait donc intéressant d'évaluer de manière approfondie l'activité biologique de cette molécule afin de mieux définir le mécanisme d'action.

***Tabella 1**. Activité in vitro des composés **1a-l** contre les lignées cellulaires K562, U937, HL60, Jurkate U266, et cytotoxicité contre les cellules mononucléaires du sang périphérique humain PBMNC + PHA.*

IC_{50} values (μM)[a]						
Inhibiteur	**K562**	**U937**	**HL60**	**Jurkat**	**U266**	**Cytotoxicité contre les cellules mononucléaires actives du sang périphérique humain (PBMNC) PBMNC + PHA**
A6730	17 ± 0.3	8 ± 0.2	5.5 ± 0.2	3.5 ± 0.2	n.d.[b]	>50
1a	4 ± 0.1	>50	>50	41 ± 1.2	>50	>50
1b	3 ± 0.1	21 ± 2.2	19 ± 2.8	5 ± 0.2	3.5 ± 0.1	14 ± 1
1c	14 ± 0.3	7 ± 0.3	12 ± 2.3	5 ± 0.1	>50	16 ± 2
1d	9 ± 0.3	>50	3.5 ± 0.1	4 ± 0.1	10 ± 0.8	41.3 ± 5
1e	3 ± 0.1	n.d.	10 ± 0.9	5 ± 0.2	18 ± 1.1	>50
1f	7 ± 0.2	7 ± 0.1	5 ± 0.1	2 ± 0.1	5 ± 0.1	8 ± 0.5
1g	8 ± 0.2	9 ± 0.4	6 ± 0.1	4 ± 0.1	9 ± 0.8	>50
1h	7 ± 0.3	4 ± 0.1	8 ± 0.2	6 ± 0.15	4 ± 0.1	50 ± 4
1i	12 ± 0.4	11 ± 0.9	24 ± 3	5 ± 0.1	8 ± 0.5	50 ± 6
1j	8.5 ± 0.3	>50	>50	6.5 ± 0.1	>50	>50
1k	3 ± 0.1	3 ± 0.1	3 ± 0.1	3.5 ± 0.1	3 ± 0.1	13 ± 1
1l	3.5 ± 0.1	8 ± 0.2	7 ± 0.3	3	3 ± 0.1	12 ± 1

[a]Les valeurs de IC_{50}(μM) correspondent à la moyenne ± déviation standard de trois expériences indépendantes.
[b]n.d. = pas déterminé.

La capacité de prédire la toxicité des molécules est très important dans la conception de nouveaux médicaments potentiels. Dans ce contexte, les molécules synthétisées **1a-l** ont été analysées par le logiciel "Molinspiration web-enabled" pour prédire les paramètres physico-chimiques importantes pour la pharmacologie de ces molécules, comme la lipophilie (Clog*P*) et l'absorption intestinale (calcul des surfaces polaires moléculaires, PSA). La surface polaire moléculaire (PSA) des composés de la série **1a-l** indique leur absorption intestinale (PSA < 140 $Å^2$). De plus, ces molécules ne sont pas capables de traverser la barrière hémato-encéphalique (PSA > 60 $Å^2$). A la lumière de ces prévisions, les composés présenteraient une bonne biodisponibilité orale avec un potentiel toxique à faible contre le système nerveux central. Ensuite, afin de rationaliser les résultats biologiques et déterminer les détails structurels qui pourraient être responsables, les études de *docking* ont été réalisées en utilisant le serveur Dock suisse Swiss Dock (http://www.swissdock.ch). La protéine considérée pour l'analyse initiale a été la structure cristallographique de la totalité de la protéine Akt déposée dans la Protein Data Bank (cod. 3O96 [6]). Les résultats de cette étude indiquent la

molécule comme le meilleur parmi celles synthétisées, mettant en évidence que la présence d'un cycle phényle dans la troisième position et l'introduction du petit groupe d'ester éthylique sur la seconde position de l'anneau quinoxaline peuvent favoriser une meilleure interaction entre le ligand et la protéine, comme illustré sur la **Figure 2**.

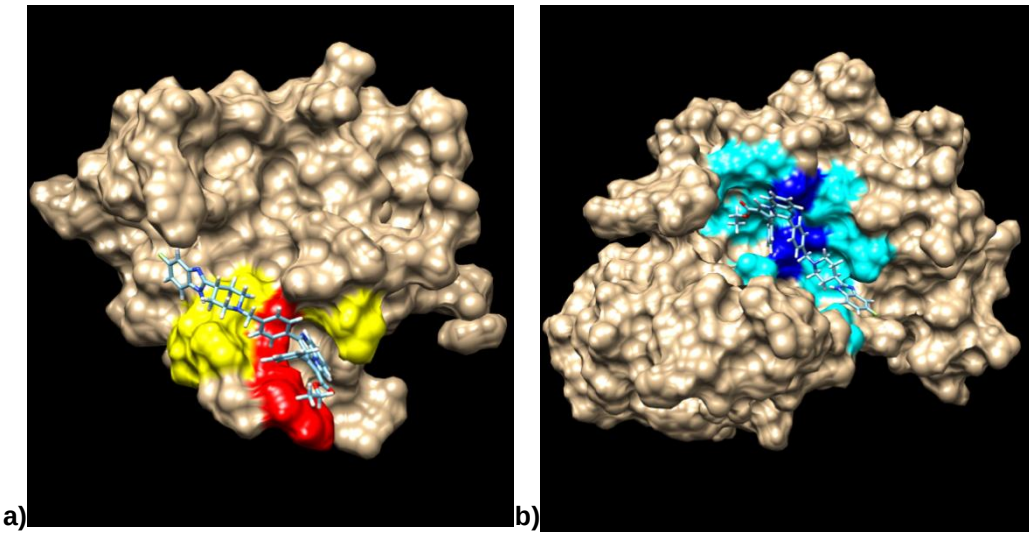


Figura 2*. Meilleure structure obtenue par le docking du composé* ***1e*** *sur la protéine entière Akt* ***(****cod. 3O96 [6]);* ***a)*** *Domaine PH: en rouge, les résidus impliqués dans au moins 10 interactions et en jaune, les résidus impliqués dans au plus 9 interactions.* ***b)*** *Domaine kinase: en bleu, les résidus impliqués dans au moins 10 interactions et en cyan, les résidus impliqués dans au plus 9 interactions.*

Ensuite, les études d'amarrage ont été menées uniquement sur le domaine PH extrapolée à partir de la structure cristallographique de PKBα (code PDB: 1UNR) [7]. Cette analyse a été effectuée afin de vérifier si notre meilleure molécule peut être considérée comme un inhibiteur allostérique potentiel de la protéine Akt ou un ligand simple PH-dépendante. En fait, la superposition de sa meilleure structure obtenue par le *docking* sur le domaine de PH de Akt et de la structure cristalline de l'inositol-(1,3,4,5) -tetrakisfosfato (IP4) dans le complexe avec le domaine PH de Akt1 (code PDB: 1UNQ [7]) montre la présence de notre ligand dans une poche séparée de le site de liaison de l'inositol. En conclusion, les résultats de *docking* indiquent que les molécules **1a-l**, plutôt que des ligands simples PH-dépendante, constituent des inhibiteurs allostériques en raison de leur capacité à interagir avec les deux domaines de la protéine Akt, qui sont les domaines PH et kinase.

Workpackage 1

Le Workpackage 2 est centré sur le CXCR4, un récepteur couplé à la protéine G qui est impliquée dans le *signaling* de facteurs différents. Dans le travail, il est décrit le comportement structurel et l'activité biologique entre CXCR4 et son ligand naturel CXCL12 [8]. L'interaction de CXCL12-CXCR4 joue un rôle important dans différentes maladies (VIH, le cancer, le syndrome WHIM, la polyarthrite rhumatoïde, la fibrose pulmonaire et le lupus). Pour cette raison, cette protéine est indiquée comme une cible thérapeutique valide [9]. Compte tenu de la puissance élevée et une spécificité élevée que les peptides peuvent avoir [10], le développement d'outils diagnostiques et thérapeutiques de nature peptidique a suscité un beaucoup d'intérêt, comme démontré par les données de la littérature récente. En particulier, les peptides intrinsèquement désordonnées (IDP) représentent actuellement un point de départ novateur et

intéressant, parce que le désordre conformationnel, définie comme l'absence de tout élément de structure secondaire régulière, telles que les hélices et les feuillets bêta, le peptide permet de subir des changements conformationnels seulement à la suite de l'interaction avec les systèmes cibles. Cette capacité de les IDPs leur donne un certain nombre d'avantages par rapport à leurs congénères structurés, telles qu'une plus grande surface d'interaction et un cinétique de liaison améliorée. En fait, ces peptides explorent dynamiquement un ensemble de conformations absolument flexibles, en adoptant le plus stable seulement après la liaison à des ligands [11]. Les IDPs peuvent, par conséquent, établir des interactions avec une spécificité élevée / faible affinité avec leurs partenaires d'interaction. À la lumière de cela, nous avons conçu deux peptides linéaires, PeptideE (YGECPCE-allyle) et PeptideK (YGECPCK- allyle), avec une teneur élevée de résidus capable de favoriser des structures désordonnées [12], en utilisant la région N-terminale de CXCL12 comme modèle, et la séquence CPC, l'introduction de la séquence CPC, décrite comme responsable de l'interaction entre le peptide CXCL12 et récepteur CXCR4 [8]. Dans ces séquences ont été ajoutées aux extrémités C-terminales des acides aminés chargés (Glu et Lys, respectivement pour la PeptideE et PeptideK) afin d'analyser l'effet de la charge sur les propriétés de liaison du peptide. L'extrémité C-terminale de ces peptides a également été modifié par un groupe ester d'allyle pour toute nouvelle fonctionnalisation des molécules. Les spectres CD ont montré l'absence d'éléments réguliers de structure secondaire et la nature désordonnée des PeptideE et PeptideK. Ce résultat a été confirmé par la faible dispersion spectrale des spectres 1D [1H] et par l'absence presque totale de signal dans les spectres 2D [1H, 1H] NOESY, qui indiquerait pour les deux peptides une plus grande flexibilité et une nature désordonnée / pas pliée. L'évaluation des déviations de déplacement chimique des Ha proton par rapport aux valeurs *random coil*, qui semblent être la plupart petites et positives, a permis de confirmer la présence de structures prolongés désordonnées. Enfin, l'état désordonné de peptides et peptide a été confirmé en outre par les patterns ROE obtenu, qui montrent des contacts séquentiels (Hα*i*–HN*i+1*) forts, typiques de peptides pas organisés structurellement.
Les résultats de l'enquête RMN de la confirment donc la flexibilité de ces peptides et, par conséquent, leur incapacité à prendre des conformations bien définies. L'étude consécutif des simulations de dynamique moléculaire (MD) (**Figure 3**), menée sur les peptides libres, a confirmé ces résultats sur la flexibilité et l'absence d'une structuration, mais a également mis en évidence la présence d'un réseau de liaisons hydrogène intramoléculaires temporaires et dynamiques de type chaîne principale-chaîne principale (MM) et la présence des interactions avec les molécules d'eau.

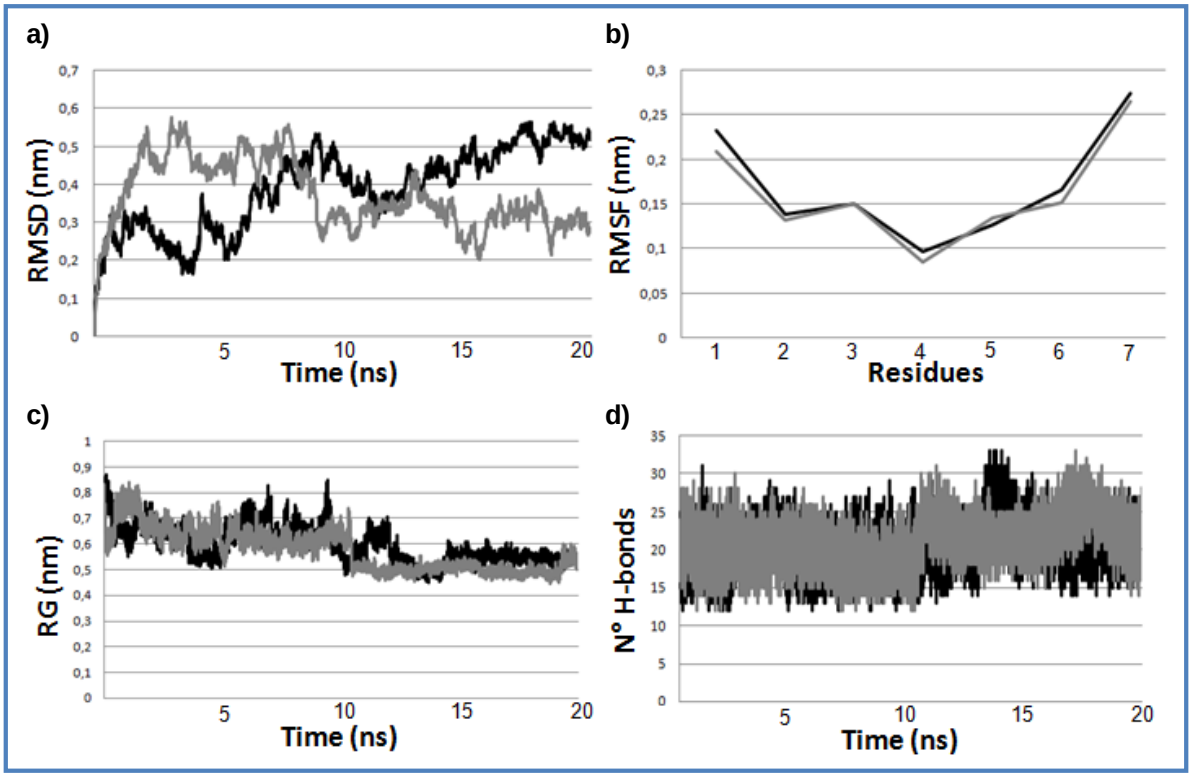


***Figure 3**. Analyse des simulations de dynamique moléculaire menée sur PeptideE (en noir) et sur PeptideK (en gris) à pH physiologique dans un solvant explicite: **a)** déviation de la racine moyenne quadratique (RMSD), **b)** les fluctuations de la racine moyenne quadratique (RMSF), **c)** rayon de rotation (gyration radius), and **d)** nombre de liaisons hydrogène.*

Par conséquent, les structures des deux peptides peuvent être représentés sous forme d'ensembles conformationnels caractérisés par des structures secondaires irréguliers et fluctuant. En fait, les peptides ne sont pas caractérisés par une conformation unique et définie, mais allant en continu d'un cluster à un autre (**Figure 4**). Par conséquent, ils ne peuvent pas être correctement décrites par une seule structure, mais par un "ensemble conformationnelle".

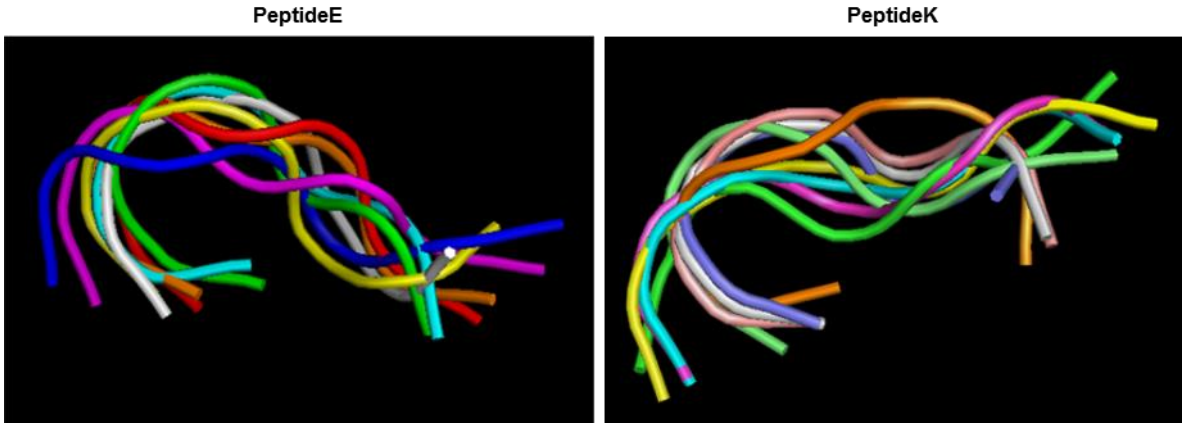


***Figura 4.** Superposition des clusters les plus peuplées obtenue pour PeptideE (8 clusters) et PeptideK (9 clusters) pendant leurs dynamiques moléculaires.*

Étant donné que l'une des difficultés majeures dans la conception d'un médicament potentiel est constitué dans son transport vers le site d'intérêt, avec tous les problèmes qui s'y rattachent (comme le franchissement des membranes cellulaires), on a pensé à fonctionnaliser l'extrémité N-terminale des peptides E et K par queues hydrophobes et donnant, donc, un caractère amphiphile à les peptides libres. Les peptides

amphiphiles sont largement décrits dans la littérature comme d'excellents candidats pour le *drug delivery* de médicaments en raison de leur aptitude à former des agrégats supramoléculaires (tels que vésicules et micelles), dû à la coexistence d'une partie hydrophile et une autre partie hydrophobe. Dans ce contexte, les peptides amphiphiles (PAs, **Figure 5**) $(C18)_2$-PeptideE et $(C18)_2$-PeptideK ont été obtenus en conjuguant deux chaînes aliphatiques de 18 atomes de carbone à l'extrémité N-terminale des deux séquences. En outre, deux autres dérivés PAs ont été synthétisés: $(C18)_2$-L-PeptideK et $(C18)_2$-L-PeptideE, caractérisé par la présence d'un *linker ethoxy* (AhOh) entre la séquence peptidique et la queue hydrophobe (**Figure 5**). Cette *linker*, le même pour les deux peptides, permet d'augmenter l'hydrophilie du peptide sans modifier la charge. De cette façon, les interactions électrostatiques intra- et intermoléculaires sont limitées. En outre, il permet d'éloigner le peptide bioactif de la coquille des agrégats hydrophobes. La longueur relativement petite de cet espaceur a été choisie pour éviter la formation d'une poche hydrophile, capable de masquer les séquences d'acides aminés [13]. Relativement faible longueur de cette entretoise a été choisie afin d'éviter la formation d'une poche hydrophile, capable de masquer les séquences d'acides aminés [13].

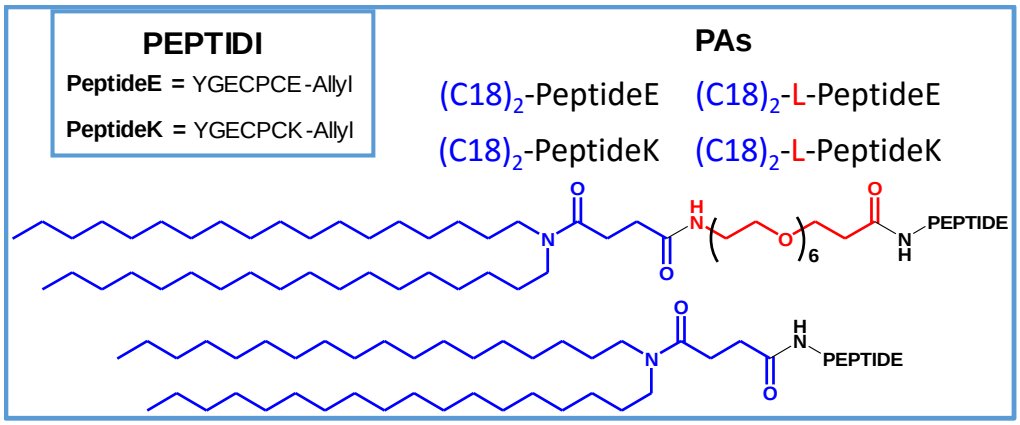


Figura 5. *PeptideE e PeptideK libres et leurs dérivés PAs. (en **bleu**, la queue hydrophobe et, en **rouge**, le linker lieur introduit). La chaîne PEG est délimitée par des parenthèses.*

Les agrégats supramoléculaires peptidiques ont été obtenus en dissolvant et en dispersant les dérivés amphiphiles du PeptideE dans un tampon phosphate 10 mM (pH 7,4) et celles du PeptideK dans un tampon TRIS 10 mM (pH 8,0). Le choix du tampon à utiliser a été imposée par les différentes propension à agréger des peptides PAs. Comme l'ont montré les études de fluorescence et DLS, tous les dérivés amphiphiles conçus sont capables d'auto-assembler pour former des agrégats avec un diamètre moyen compris entre 80 et 110 nm, indépendamment de la charge à l'extrémité C-terminale et la présence du *linker*.

Les solutions de chaque agrégat montrent une répartition monomodale
L'ensemble de chaque solution d'AP présentent une distribution monomodale, qui indique la présence d'une seule population des agrégats. Dans le Tableau 2 sont rapportées les données obtenues par l'analyse de fluorescence et DLS de tous les composés: concentration critique micellaire (**CMC**), rayon hydrodynamique ($\boldsymbol{R}_H$), les coefficients de diffusion (***D***), et un indice de polydispersité (**PDI**).

Tabella 2*. Paramètres structuraux obtenus pour les agrégats par des mesures de fluorescence (concentration micellaire critique CMC) et mesures DLS (rayon hydrodynamique R_H, coefficients de diffusion D, et indices de polydispersité PDI).*

Peptide	CMC (mol kg^{-1})	R_H (nm)	D x 10-12 (m^2s^{-1})	PDI
$(C18)_2$-PeptideK	7.0×10^{-6}	80 ± 27	3.1 ± 1.0	0.270
$(C18)_2$-L-PeptideK	1.9×10^{-5}	82 ± 34	3.1 ± 1.5	0.182
$(C18)_2$-PeptideE	3.9×10^{-5}	112 ± 44	1.9 ± 0.7	0.264
$(C18)_2$-L-PeptideE	3.2×10^{-5}	110 ± 39	2.4 ± 0.5	0.226

Les spectres CD obtenus pour les dérivés Pas (acquis à une concentration supérieure à la concentration micellaire critique) montrent que la présence du linker ne modifie pas significativement la structure de la partie peptidique. En outre, ces spectres ont montré que la présence d'un résidu de Lys par rapport au résidu Glu à l'extrémité C-terminale peut conduire à une réorganisation de la structure secondaire avec un *folding* partiel de la partie peptidique, confirmant l'importance de la charge dans ces séquences peptidiques. Les composés synthétisés par nous sont les premiers peptides amphiphiles contenant la séquence CPC et capable d'agréger et de former des structures supramoléculaires avec une activité *in vitro*. En fait, les tests biologiques réalisés sur les mêmes cellules et dans les mêmes concentrations utilisées pour la molécule de référence AMD3100, ont montré que, malgré tous les peptides ne présentent pas une affinité de liaison évident pour CXCR4 (données non illustrées), ils sont capables d'induire une production de cAMP supérieure à celle de AMD3100 et d'une manière dépendante de la dose. Ces données indiquent une activité inhibitrice significative des peptides synthétisés vers CXCR4. Les valeurs de cAMP de tous les peptides et celles de AMD3100 (molécule de référence) sont résumés dans la **Figure 6**, où toutes les valeurs sont exprimées comme concentration en pourcentage par rapport à deux concentrations (1μM et 10μM). Le peptide CXCL12, le ligand naturel de CXCR4, a été examinée dans cette étude comme témoin.

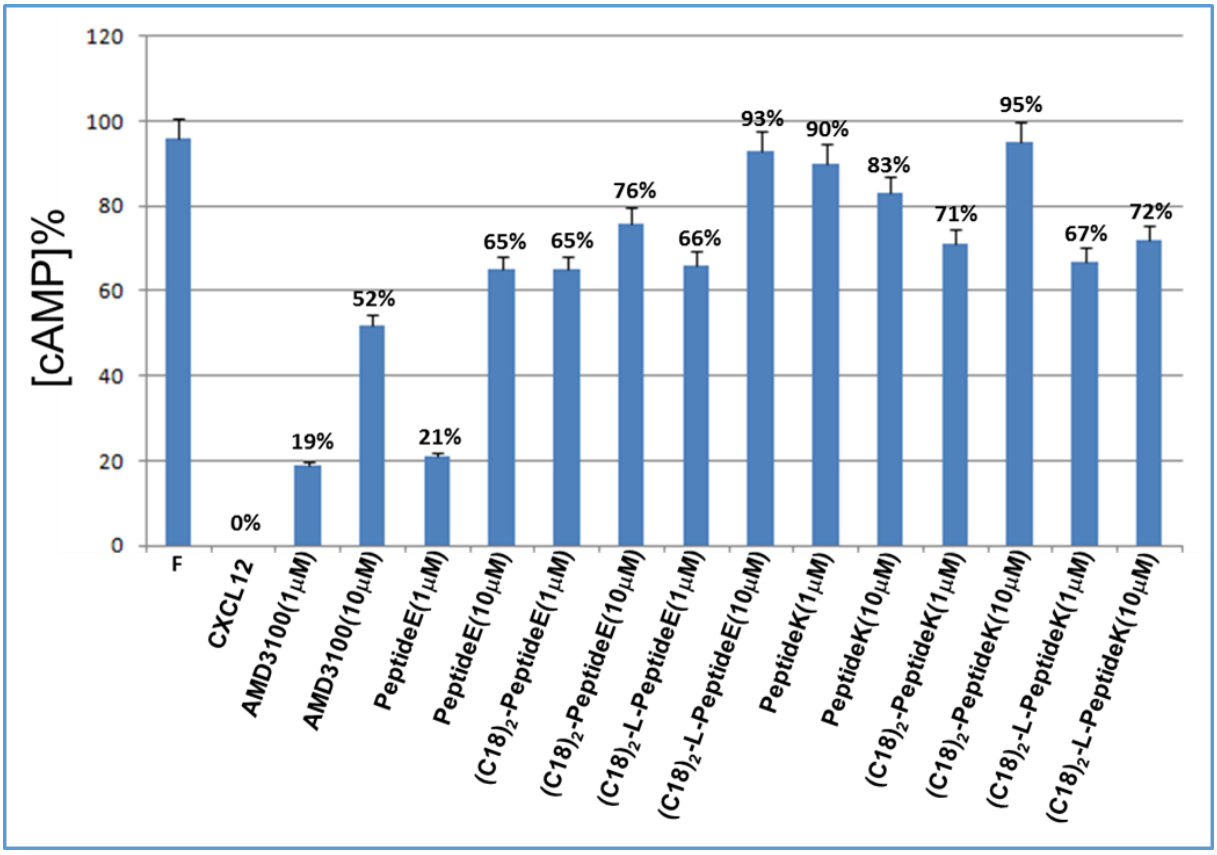


***Figure 6**. Comparaison des modulations de cAMP évalués pour Fosfokolin (F), CXCL12, PeptideE, PeptideK, $(C18)_2$-PeptideE, $(C18)_2$-PeptideK, $(C18)_2$-L-PeptideE, $(C18)_2$-L-PeptideK et pour la molécule de référence AMD3100, aux concentrations de 1 μM et 10 μM.*

Bien que les résultats doivent être considérés comme préliminaires et même si une analyse plus approfondie est nécessaire, nous pouvons certainement dire que ces résultats sont encourageants et montrent la validité de l'approche multidisciplinaire à la base de de cette investigation scientifique. L'étape suivante sera dans la direction d'apporter d'autres modifications chimiques, en utilisant la réactivité du groupe ester d'allyle, pour obtenir de nouveaux composés supramoléculaires de nature peptidique, capables d'être encapsulés dans les cellules et, par conséquent, de transmettre le peptide bioactif et / ou des sondes pour thérapies médicales individuelles et / ou pour des investigations en imagerie par résonance magnétique (IRM).

1. Paul S.M. *et al*., Nat.Rev. Drug Discov., 2010, 9(3), 203-214;
2. Yu Y. et al., Plos One, 2015, 10(10), e0140479;
3. Zabel B. A. *et al.*, J.Immunol., 2009, 183, 3204–3211;
4. Ma Q. *et al.,* Proc. Natl. Acad. Sci. USA, 1998, 95, 9448–9453;
5. Desplat V. et al., J.Enzym.Inhib. Med. Ch., 2011, 26(5), 657-667;
6. Wu-I W. et al., *Plos One*, 2010, 5(9), e12913;
7. Milburn CC et al, Biochem. J., 2003, 375, 531-538;
8. Costantini S. *et al.*, J.Pept. Sci., 2014, 20 (4), 270-278;
9. Tamamura H. e Fujii N., *Expert Opin. Ther. Tar.*, 2005, 9(6), 1267–1282;
10. Accardo A. *et al.*, Mol. BioSyst., 2013, 9, 1401-1410;
11. Mittal J., J. Phys. Chem. B, 2013, 117 (1), 118-124;
12. Uversky V. N., *Int.J.Biochem. Cell B.*, 2011, 43(8), 1090-1103;
13. Matsumura Y. & Maeda H., CancerRes., 1986, 46, 6387-6392.